Høyfrekvens finans og markedets mikrostruktur på Oslo Børs

Arne Danielsen

Masteroppgave i statistikk
Finansteori og forsikringsmatematikk

Matematisk institutt
Universitetet i Bergen
18. september 2009

Takk

Jeg ønsker først og fremst å takke min veileder Hans A. Karlsen for god veiledning og mange nyttige innspill under arbeidet med masteroppgaven, og for å ha gitt meg stor frihet i utformingen av oppgaven. Vil også takke alle medstudenter ved Kroepeliens hus, samt familie som har bidratt med korrekturlesing.

Innhold

1 Innledning

Høyfrekvens finans, som er tema for denne oppgaven, tar for seg studiet av høyfrekvente finansielle data. I denne oppgaven har vi sett på data med den høyest mulige frekvensen, nemlig data fra hver eneste handel.

De høyfrekvente dataene forteller oss tidspunktet for en handel og hva prisen var ved handelen. Differensierer vi pris og tidspunkt, får vi prisendringen y_i og tidsavstanden τ_i som vi studerer nærmere. Disse dataene viser seg å ha flere spesielle egenskaper. Bl.a. viser det seg at prisendringene kun tar noen få verdier, se 3.1. Det er en utfordring å finne en modell som passer til slike data. I Bauwens *et al.* (2008, s. 2) nevnes det to tilnærminger til dette problemet. Den ene løsningen er å se på dataene som en brownsk bevegelse med støy på, som vi går inn på i kapittel 6. Den andre løsningen er å se på dataene som strukturell informasjon. Vi har i denne oppgaven sett mest på den siste tilnærmingen og studert ACM-ACD-modellen foreslått av Russell og Engle (2005), se kapittel 4. I denne modellen tenker man seg at y_{i-1} har en effekt på τ_i, og at τ_i har en effekt på y_i. I denne oppgaven kommer vi frem til at kun det siste er tilfelle. Ved å lage et kryssdiagram finner vi ingen sammenheng mellom y_{i-1} og τ_i, og ved estimering av ACM-ACD-modellen kommer vi frem til at y_{i-1} ikke forklarer τ_i, se 3.5. Vi har derfor i denne oppgaven forenklet ACM-ACD-modellen noe. Russell og Engle (2005) hevder at det er nummerisk effektivt å maksimere de to delene i ACM-ACD-modellen simultant. Når vi utelater effekten av y_{i-1} på τ_i, kommer vi i denne oppgaven frem til at det er nummerisk effektivt å maksimere de to delene hver for seg.

Modellen estimeres ved sannsynlighetsmaksimeringsestimering, og vi får da også ut estimerte standardavvik. Det er i Russell og Engle (2005) ikke diskutert hvor gode disse estimatene er. For å finne svar på dette sammenligner vi estimatene med estimater vi får ut fra et gjentaksforsøk, se 4.5.

Et annet problem ved estimering av modellen er at vi ikke kan være sikre på om algoritmen finner et globalt maksimum. Diagnostiske tester gir oss et svar på dette spørsmålet, samtidig som det gir oss et svar på hvor bra modellen passer til dataene. Men verken Russell og Engle (2005) eller Bauwens *et al.* (2008, kap. 8) kommer frem til

gode metoder for å vurdere modellens tilpasning. Problemet er i denne oppgaven løst ved å estimere modellen med de reelle dataene, simulere modellen med parameterverdiene vi får ut og deretter sammenligne de reelle og de simulerte dataene. Vi observerer da at marginale fordelinger til y og τ, autokorrelasjon og sammenheng mellom τ_i og y_i er de samme i reelle og simulerte data, se 5.4.2.

En viktig problemstilling er også hvor store datamengder modellen krever. Gjennom simuleringseksperiment kommer vi frem til at vi må ha over 10 000 data for at modellen skal gi gode resultater, se 4.5.

Videre tolker vi i oppgaven resultatene vi får ut fra modellen, se 5.5. Modellen gir oss innsikt i markedets mikrostruktur, og en slik innsikt gir oss en forståelse av hvordan markedet fungerer. Ved å tilpasse modellen til ulike aksjer, og særlig ved å studere aksjer omsatt på ulike børser, vil man kunne se forskjeller. Høsten 2008 ble det forbudt med kort salg i enkelte aksjer på Oslo Børs, og i følge Russell og Engle (2005) skal dette gi utslag i modellen. Teorien sier nemlig at dersom man ikke har mulighet for kort salg, vil store tidsavstander være et signal om at prisen er på vei nedover. Dette kan vi teste ved å teste for symmetri i ACM-ACD-modellen.

Et interessant spørsmål i forbindelse med høyfrekvente data er om det er mulig å bruke dataene til å tjene penger med. Her kommer temaet triangelarbitrasje inn. Dersom vi har to valutakurser USD/SEK og USD/NOK, vil valutakursen NOK/SEK være gitt som forholdet mellom de to første. Holder ikke dette, vil vi ha en arbitrasjemulighet. For å modellere de tre tidsrekkene USD/SEK, USD/NOK og NOK/SEK kan vi bruke en kointegrasjonsmodell. Trapletti *et al.* (2002) viser at vi har kointegrasjon i et datasett med valutakursene USD/DEM, USD/JPY og DEM/JPY, mens vi i denne oppgaven påviser kointegrasjon i valutakursene USD/SEK, USD/NOK og NOK/SEK, se 7.2. Videre forsøker vi å tolke resultatene fra kointegrasjonsmodellen, og vi sammenligner kointegrasjonsmodellens prediksjonsevne med en VAR-modell.

Programvare

For å simulere ACM-ACD-modellen har vi skrevet et skript i MATLAB, se B.1, basert på algoritmen i 4.4.1. Programkoden for estimering av ACM-ACD-modellen har vi skrevet i R, se B.2. Koden baserer seg på algoritmen beskrevet i 4.3.2 og benytter i tillegg R-pakken maxLik.

For å tilpasse kointegrasjonsmodellen i kapittel 7, har vi benyttet R-pakkene urca og vars som er nærmere beskrevet i Pfaff (2008, s. 163).

2
Handel av aksjer

2.1 Handel på Oslo Børs

Frem til 1988 ble handelen på Oslo Børs utført ved en oppropsordning. Da ble denne ordningen erstattet av et elektronisk handelsstøttesystem. Dette systemet ble erstattet av et nytt elektronisk system i 1999, som blant annet gjorde det mulig for meglerne å flytte ut av børsbygningen. I tillegg ble handel over internett mulig, noe som førte til en sterk økning av handel i følge OSE (2007a, s. 8-9).

2.1.1 Handelsdagen

Oslo Børs er åpen mandag til fredag kl. 09.00-17.30 og handelsdagen inndeles som følger:

Inndeling av børsdagen:	
kl. 08.15-09.00	Før-børs perioden
kl. 09.00	Åpningsauksjon
kl. 09.00-17.20	Kontinuerlig handel
kl. 17.20-17.30	Sluttauksjon

Fra kl. 08.15 kan meglerne begynne å registrere order de har fått fra kundene. Klokken 09.00 åpner børsen, og man starter da med en åpningsauksjon. Meglerne kan se hvilke ordrer de andre meglerne har lagt inn, og de kjøps- og salgsordrer som passer sammen, fører til en handel. Fra kl. 09.00-17.20 gjennomføres kontinuerlig handel. I denne perioden legger meglerne inn ordrer med en gang de blir mottatt, og disse blir utført med engang dersom det er kjøps- og salgsordrer som passer sammen. Kl. 17.20 stanses all handel, og alle som ønsker å handle, har 2-3 minutter på seg til å legge inn ordrer. Mellom kl. 17.23 og 17.28 gjennomføres en sluttauksjon.

I 2006 var gjennomsnittlig dagsomsetning 10,3 milliarder kroner per dag, og gjennomsnittlig antall transaksjoner var 35 200 i følge OSE (2007b, s. 5).

2.1.2 Meglerens rolle

En investor kan ikke legge inn ordrer direkte på Oslo Børs. Det er kun meglerhus som har fått konsesjon fra Kredittilsynet som kan gjøre dette. Ved utgangen av 2006 var det, i følge OSE (2007b, s. 8), 52 slike meglerhus, og 28 av disse var utenlandske. En av fordelene med meglere som mellomledd, er at det sikrer at partene i handelen har dekning for den innlagte ordren, dvs. at man har finansiell dekning dersom man ønsker å handle, og at man ikke forsøker å selge et verdipapir man ikke har. Meglerne har også en rolle når det gjelder å utføre ordrer mellom egne kunder og rapportere dette til børsen. Meglerhusene driver også med egenhandel, dvs. kjøp og salg for meglerhusets egne penger og dette omtales i Grøtte (2006, s. 427).

I dag blir såkalte nettmeglere stadig mer brukt. Ved utgangen av 2005 utgjorde slik handel ca. 20 prosent av den totale omsetning og nær 45 prosent av alle transaksjoner i følge OSE (2006, s. 28). Ved bruk av nettmegler legger man inn en ordre fra sin egen datamaskin, som går gjennom meglerens system og inn i børsens handelssystem. Et alternativ til nettmegler er å kjøpe aksjer gjennom banker eller rene meglerhus. Fordelen med dette er at man får investeringsråd og anbefalinger, men kurtasjen, dvs. det man må betale i meglerprovisjon, er betydelig høyere i dette tilfellet enn ved å bruke en nettmegler.

2.1.3 Handel i aksjer

Når man ønsker å legge inn en ordre, kan man velge mellom limit-ordre og best-ordre. Ved en limit-ordre setter man en øvre grense for hva man er villig til å betale for aksjen. Man får da ikke kjøpt før noen er villig til å selge på denne kursen eller lavere kurs, og tilsvarende gjelder for salgsordre. Ved en best-ordre er det opp til megler å handle på den mest fordelaktige kursen. Ved best-ordre har man ikke noen garanti for hva prisen vil bli, men sannsynligheten for at ordren gjennomføres er større. Man må også avgjøre hvor lenge ordren skal være gyldig, og hvor stort volum man ønsker å handle.

Oslo Børs fungerer som et ordre-drevet marked hvor bud og tilbud tastes inn i ordreboken og automatisk møtes dersom pris, volum og andre betingelser er oppfylt, og omtales i OSE (2007b, s. 20). Illustrasjonen nedenfor viser et eksempel på ordreboken for StatoilHydro.

Akkumulert ordrebok for STL på Oslo Børs

Kjøp			Salg		
Ordre	Dybde	Bud	Tilbud	Dybde	Ordre
5	14 600	151.00	151.10	14 314	6
3	21 150	150.90	151.20	5 100	3
3	1 750	150.80	151.30	14 000	2
1	1 000	150.70	151.40	6 300	5
2	3 700	150.60	151.50	5 000	1
166	464 820	Total kjøpere	Total selgere	478 491	220

Figur 2.1: Utdrag fra ordrebok for StatoilHydro 6. august 2008 hentet fra Netfonds.no

Ut fra denne ordreboken ser man at den høyeste prisen noen er villig til å kjøpe for er 151,00 kr, mens laveste pris noen er villig til å selge for er 151,10 kr. Ønsker man å kjøpe en aksje, kan man enten legge inn et bud på 151,10 kr eller man kan vente til selgerne evt. setter ned prisen. Dersom noen har likt bud, får den som la inn budet først, handle. Legger man inn et bud på 151,00 kr, vil man i dette tilfellet ha fem kunder foran i køen som får handle først.

Ønsker man å selge, kan man enten legge inn et tilbud på 151,00 kr og få solgt alle aksjene med engang, så lenge antall aksjer man ønsker å selge ikke overstiger de 14 600 aksjene som ønskes kjøpt til denne prisen. Alternativt kan man vente til kjøperne evt. hever sitt bud. Tilsvarende som ved kjøp av en aksje, vil det ved likt bud være den som legger inn budet først som får selge.

Kun handel i hele børsposter havner i ordreboken. En børspost tilsvarer aksjer for ca. 10 000 kr. Man kan også handle i såkalte odd-lots, som er et volum som ikke tilsvarer en hel aksjepost. Odd-lots er vanskeligere å omsette fordi det må finnes en motpart som vil handle i odd-lots til samme pris.

Vi merker oss at i ordreboken på forrige side, er differansen mellom de ulike budene 0,10 kr. Dette har sammenheng med aksjens tickstørrelse. Tickstørrelse defineres som følger:

Definisjon 2.1.1: Tickstørrelse

Dette er den minste prisendring som kan tastes inn i handelssystemet. Størrelsen på en aksjes tickstørrelse vil være avhengig av prisen på aksjen.

For aksjer notert på hovedindeksen OBX i 2007 var tickstørrelse i følge OSE (2007b, s. 21):

Aksjepris(NOK)	Tickstørrelse(NOK)
<15,00	0,01
15,00–49,95	0,05
50,00–99,90	0,10
100,00–249,75	0,25
250,00–499,50	0,50
>500	1,00

2.2 Handel på andre børser

Børsene i verden er organisert på ulike måter. Reglene på en børs blir utformet slik at handelen blir mest mulig likvid. I Grøtte (2006, s. 396) blir likviditet definert som et mål på hvor lett det er å få omsatt et verdipapir.

2.2.1 NASDAQ

I likhet med Oslo Børs er NASDAQ (National Association of Securities Dealers Automated Quotation System) en elektronisk børs. En forskjell som omtales i Grøtte (2006, s. 433), er at på NASDAQ fungerer meglerne som såkalte market makers. Dette innebærer at meglerne har plikt til å sette en kurs som kundene kan handle for. I likhet med Oslo Børs legges ordrene inn i en ordrebok som er tilgjengelig for investorene. Aksjene på NASDAQ har tickstørrelse som er langt mindre enn ved Oslo Børs, se Bessembinder (2003).

2.2.2 NYSE

Frem til 2006 gikk ordrene på NYSE (New York Stock Exchange) gjennom spesialister. Hver aksje hadde en spesialist, som var en person på børsen. Spesialistens oppgave var å matche ordrer, og kun spesialisten hadde full tilgang til ordreboken. Spesialisten måtte til enhver tid stille kurser. Ved store kjøpsordrer måtte spesialisten selv selge, noe som drev kursen opp, og spesialisten måtte derfor få kursen ned i etterkant for å prøve å kjøpe tilbake billigere. Tilsvarene måtte han kjøpe ved store salgsordrer, noe som drev kursen ned, og spesialisten måtte få kursen opp i etterkant for å kunne selge høyere. Hensikten med dette systemet var i følge Grøtte (2006, s. 434), å få mindre volatilitet.

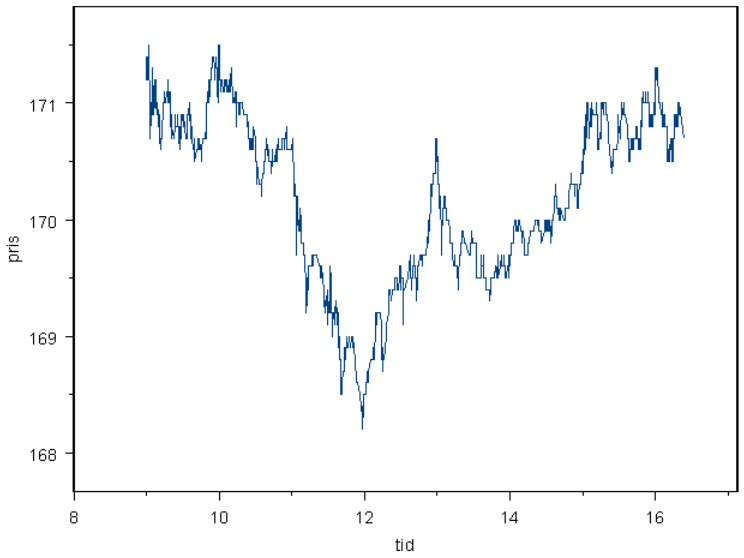
Fra 2006 har NYSE fungert som et hybridmarked, omtalt i Hendershott og Moulton (2007). Med dette menes at man kan enten legge inn en ordre i det automatiske systemet eller sende inn en ordre til spesialisten. En av hensiktene med denne omleggingen var å få et raskere system, ettersom handelen går fortere gjennom et automatisk system. Dette omtales i NYSE (2006, kap. 2). I likhet med NASDAQ har aksjene på NYSE langt mindre tickstørrelse enn ved Oslo Børs, se Bessembinder (2003).

3
Høyfrekvente data fra Oslo Børs

I dette kapittelet studerer vi høyfrekvente data fra Oslo Børs, og vi har tatt for oss data fra aksjen StatoilHydro. I datasettet finner vi observasjoner fra hver eneste handel, og hver handel er registrert med et tidspunkt og en pris. Som det fremkommer i dette kapittelet, har disse dataene flere spesielle egenskaper.

3.1 Diskrete prisendringer

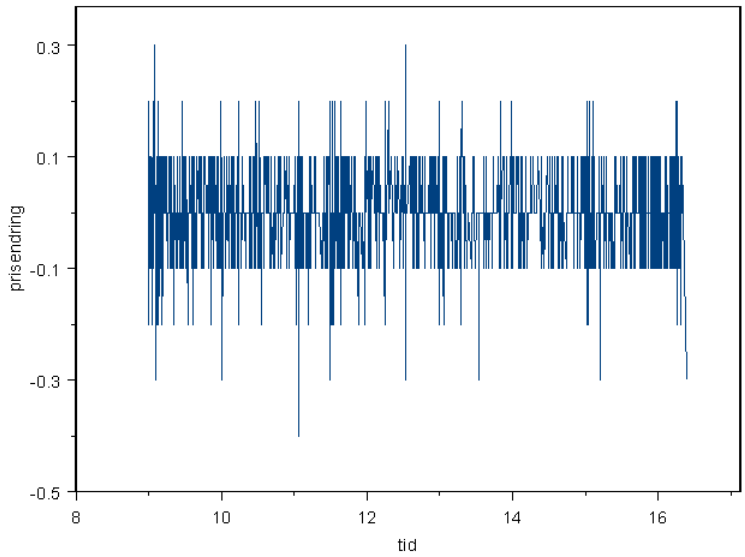
Figurene nedenfor viser aksjeprisens utvikling i løpet av en dag, plottet for hver eneste handel.



Figur 3.1: Aksjepris i kroner StatoilHydro plottet mot tidspunkt på dagen 16. april 2008

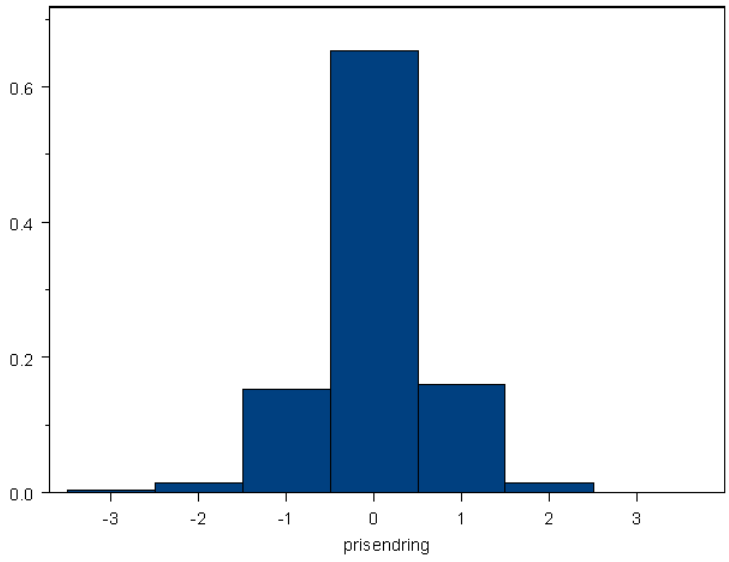
Den neste figuren viser prisendringen fra en handel til den neste. Ut fra denne figuren ser man at prisendringen kun tar noen få verdier. Årsaken til dette er at aksjeprisen

kun kan endre seg i hele tick, se definisjon 2.1.1 på side 11. I dette tilfellet ser man ut fra figuren at tickstørrelse er lik 0,1.



Figur 3.2: Prisendringen i kroner StatoilHydro plottet mot tidspunkt på dagen 16. april 2008

Figuren og tabellen under viser den empiriske fordelingen til prisendringene y. Vi ser at dette er en symmetrisk fordeling, og at over 60 % av handlene fører ikke til noen prisendring. Antall observasjoner n er i dette tilfellet lik 2060.



Figur 3.3: Empirisk fordeling til prisendringene i tickstørrelse StatoilHydro 16. april 2008. Tickstørrelse er her lik 0,1.

k	-3	-2	-1	0	1	2	3
$\mathbb{P}(y = k)$	0,003	0,015	0,153	0,652	0,161	0,015	0,001

3.2 Autokorrelasjon og bidask-effekt

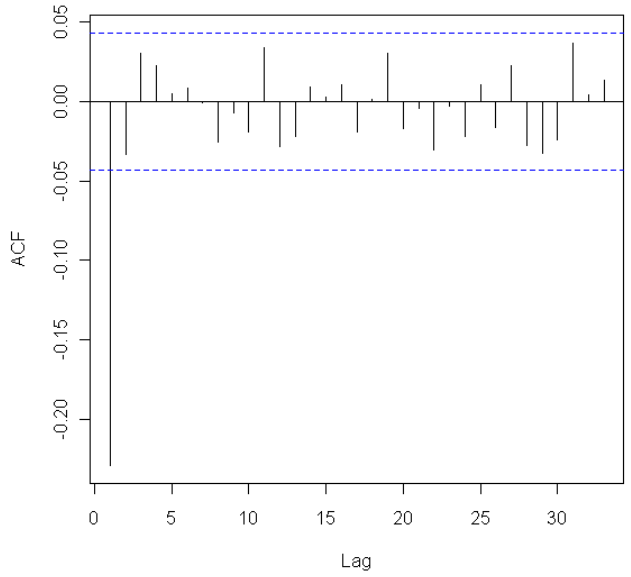
Autokorrelasjonsfunksjonen(ACF) beskrives i Tsay (2005, s. 26-27) og er gitt som

$$\rho_l = \frac{\mathrm{Cov}(y_i, y_{i-l})}{\sqrt{\mathrm{Var}(y_i)\,\mathrm{Var}(y_{i-l})}},$$

der l er antall lag. Autokorrelasjonsfunksjonen estimeres som følger:

$$\hat{\rho}_l = \frac{\frac{1}{n-l-1}\sum_{i=l+1}^{n}(y_i - \bar{y})(y_{i-l} - \bar{y})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\frac{1}{n-l-1}\sum_{i=l+1}^{n}(y_{i-l} - \bar{y})^2}}.$$

Dersom man regner ut autokorrelasjonen til prisendringen $y_i = p_i - p_{i-1}$, der p_i er prisen ved handel nummer i og p_{i-1} er prisen ved handel nummer $i-1$, vil man se at prisendringen har negativ lag(1)-korrelasjon, som figuren nedenfor viser.



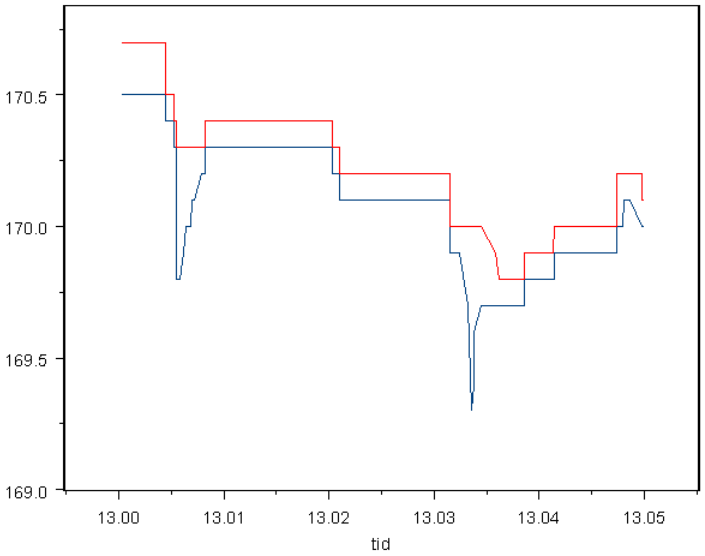
Figur 3.4: Autokorrelasjon til prisendringene StatoilHydro 16. april 2008

Dette ser man også dersom man regner ut hvilken retning den neste prisendringen har etter en henholdsvis positiv, null, og negativ prisendring. Tabellen på neste viser at dersom vi har hatt en positiv prisendring, er sannsynligheten for at vi vil ha negativ prisendring ved neste handel mye større enn at vi vil ha en ny positiv prisendring. Har vi hatt en negativ prisendring, er sannsynligheten for en positiv prisendring ved neste handel større enn en negativ prisendring. For å regne ut tabellen nedenfor har vi brukt data fra perioden 6. mars til 3. juni for StatoilHydro-aksjen. Dette er et datasett med antall observasjoner n lik 107 379. Vi observerer at vi har symmetri i matrisen.

i-1\ i	-	0	+
−	**0,11**	0,53	**0,36**
0	0,18	0,64	0,18
+	**0,36**	0,53	**0,11**

Tabell 3.1: Virkning av y_{i-1} på y_i

Vi ønsker å finne ut årsaken til autokorrelasjonen i dataene, og vi ser på en forklaring som er beskrevet i Tsay (2005, s. 210-212), kalt bidask-effekten. Fenomenet knyttes til måten aksjer handles på. Øverst i ordreboken, se figur 2.1 på side 10, finner man det til enhver tid høyeste kjøpsbudet(bid) og laveste salgsbudet(ask). Når disse verdiene er like, skjer det en handel, og ordreboken oppdateres. Ordreboken oppdateres også når det legges inn nye bud. Dette gir oss to tidsrekker. Figuren nedenfor viser laveste salgsbud(ask) og høyeste kjøpsbud(bid) til enhver tid i et kortere tidsintervall.



Figur 3.5: Bid (blå kurve) og ask (rød kurve) i kroner StatoilHydro 16. april 2008 i tidsrommet 13:00-13:05

Så lenge det er en forskjell mellom det kjøper vil betale og det selger vil ha for aksjen, blir det ingen handel, enten må kjøper heve sitt bud eller selger må senke sitt krav. Vi kan tenke oss at den virkelige aksjeprisen ligger midt mellom kjøper og selger. Den observerte aksjeprisen tilfredsstiller da ligningen

$$p_i = p_i^* + q_i \frac{s}{2},$$

der

- p_i er den observerte aksjeprisen,
- p_i^* er den virkelige aksjeprisen,
- q_i er en stokastisk variabel som tar verdiene -1 og 1 med samme sannsynlighet og
- s er en konstant differanse mellom selgers krav og kjøpers bud.

Dersom vi antar at det ikke er noen endring i den virkelige aksjeprisen p_i^* kan vi skrive prisendringen som

$$\Delta p_i = (q_i - q_{i-1})\frac{s}{2}.$$

Vi har da at

$$\text{Var}(\Delta p_i) = \frac{s^2}{2}.$$

Videre har vi at

$$\begin{aligned}\text{Cov}(\Delta p_i, \Delta p_{i-j}) &= \mathbb{E}(\Delta p_i \Delta p_{i-j}) \\ &= \left(\frac{s}{2}\right)^2 \mathbb{E}((q_i - q_{i-1})(q_{i-j} - q_{i-j-1})),\end{aligned}$$

som gir oss at

$$\text{Cov}(\Delta p_i, \Delta p_{i-j}) = \begin{cases} -\frac{s^2}{4} & j = 1 \\ 0 & j > 1. \end{cases}$$

Autokorrelasjonsfunksjonen til Δp_i er da

$$\rho_j(\Delta p_i) = \begin{cases} -0{,}5 & j = 1 \\ 0 & j > 1. \end{cases}$$

Den negative lag(1)-korrelasjonen kan forståes intuitivt. Dersom vi har observert en pris lik $p_i^* + \frac{s}{2}$, må prisen ved neste handel være den samme eller lavere. Tilsvarende dersom vi har observert pris lik $p_i^* - \frac{s}{2}$, må neste pris være enten den samme eller høyere.

Ovenfor har vi antatt at den virkelige prisen p_i^* ikke endrer seg. Det er mer realistisk å anta at p_i^* følger en tilfeldig gang, dvs. at

$$\Delta p_i^* = p_i^* - p_{i-1}^* = \epsilon_i,$$

som danner en følge av uavhengige og identisk fordelte stokastiske variable med forventning null og varians σ^2. I tillegg er ϵ_i uavhengig av q_i. Vi får nå at

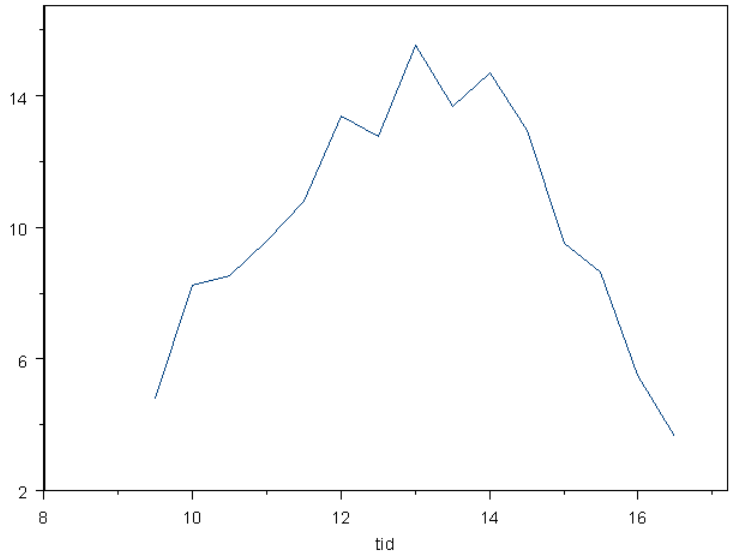
$$\begin{aligned}\text{Var}(\Delta p_i) &= \frac{s^2}{2} + \sigma^2 \text{ og} \\ \rho_1(\Delta p_i) &= \frac{-s^2/4}{s^2/2 + \sigma^2} \leq 0.\end{aligned}$$

Dette fører til at lag(1)-korrelasjonen blir noe mindre, men den vil fortsatt være negativ.

Det er viktig å merke seg at dersom den virkelige prisen tar kontinuerlige verdier og den observerte prisen tar diskrete verdier holder ikke forklaringen ovenfor helt. I kapittel 6 kommer vi nemlig frem til at dersom den virkelige prisen tar kontinuerlige verdier og den observerte prisen er avrundinger til diskrete verdier, vil dette lage negativ lag(1)-korrelasjon i dataene. Det kan dermed diskuteres om man egentlig har bidask-effekt i dataene.

3.3 Inhomogen tidsrekke og daglig periode

Vi ser i dette avsnittet på tidsavstandene $\tau_i = t_i - t_{i-1}$, der t_i er tidspunktet for handel nummer i, og t_{i-1} er tidspunktet for handel nummer $i-1$. Avstanden i tid mellom handlene τ_i er stokastisk, og vi får dermed en inhomogen tidsrekke. Med inhomogen tidsrekke mener man at avstanden i tid mellom observasjonene ikke er konstant. Den forventede tidsavstanden til neste handel vil også variere i løpet av dagen. Dette betyr at tidsavstandsprosessen heller ikke er stasjonær. I starten og slutten av dagen vil det være høyest aktivitet, og tidsavstandene blir minst. Rundt lunsjtid vil det være lavere aktivitet, og man kan få lange perioder uten handel. Dette vises ut fra figuren nedenfor.



Figur 3.6: Gjennomsnittlig tidsavstand i sekunder beregnet ut fra aksjepriser StatoilHydro 1.-15. april 2008

I modeller der tidsavstanden inngår ønsker vi å fjerne denne sesongen. Vi betrakter derfor den justerte tidsavstanden ved å utføre en splineinterpolasjon av kurven ovenfor, og vi deler alle tidsavstander på tilhørende splineverdi, se 5.1.2. Dette betyr at vi har en ikke-lineær transformasjon av tidsskalaen, og dersom vi plotter kurven i figur 3.6 etter transformeringen vil vi få en rett linje. Dette kan også oppfattes som en preprosessering. I Tsay (2005, s. 225-227) løser man problemet på en mer komplisert måte ved bruke kvadratiske funksjoner og indikatorvariabler.

3.4 Klynger i tidsavstander etter sesongjustering

Etter at vi har fjernet sesongen i tidsavstandene, vil det være rimelig å anta at tidsavstandene har konstant intensitet og er uavhenige eksponentialfordelt. Tabellen øverst på neste side viser imidlertid at dette ikke stemmer. Tidsavstandene er her justert for sesong og skalert slik at forventet tidsavstand er lik 1. Vi har i fortsettelsen valgt å kalle de sesongjusterte tidsavstandene τ_i.

	$\tau_i \leq 0{,}28$	$0{,}28 < \tau_i \leq 0{,}85$	$\tau_i > 0{,}85$
$\tau_{i-1} \leq 0{,}28$	**0,41**	0,32	0,27
$0{,}28 < \tau_{i-1} \leq 0{,}85$	0,33	0,34	0,33
$\tau_{i-1} > 0{,}85$	0,26	0,31	**0,43**

Tabell 3.2: Virkning av τ_{i-1} på τ_i

Vi har klynger i de sesongjusterte tidsavstandene. Små tidsavstander blir fulgt av små tidsavstander og store tidsavstander blir fulgt av store tidsavstander. Denne strukturen minner om den man finner i volatilitet, og tidsavstandene modelleres derfor med modeller svært lik volatilitetsmodeller, se 4.1.

3.5 Sammenheng mellom justerte tidsavstander og prisendringer

Fra tabellene nedenfor ser det ut som sannsynligheten for prisendring øker noe når det har gått lang tid siden forrige handel, mens en prisendring ikke ser ut til å ha noen betydning for forventet tid til neste handel. Dvs. at variabelen τ_i påvirker y_i, mens y_{i-1} ikke påvirker τ_i. Vi ser derfor ut til å ha funnet en årsakssammenheng, også kalt kausalitet.

	-	0	+	+/-
$\tau_i \leq 0{,}28$	0,17	0,66	0,17	**0,34**
$0{,}28 < \tau_i \leq 0{,}85$	0,20	0,60	0,20	**0,40**
$\tau_i > 0{,}85$	0,23	0,54	0,23	**0,46**

Tabell 3.3: Virkning av τ_i på y_i

Vi har her brukt et datasett på 107 379 data. For å kontrollere om vi har signifikans kan vi bruke en kjikvadrat-observator for å teste for homogenitet som foreslått i Walpole *et al.* (2002, s- 340-341). Med homogenitet mener vi her at sannsynligheten for prisendring er lik i de tre gruppene ($\tau_i \leq 0{,}28$, $0{,}28 < \tau_i \leq 0{,}85$ og $\tau_i > 0{,}85$). Siden vi i dette tilfellet har så store datasett, får vi p-verdier nær null, og vi forkaster hypotese om homogenitet.

	$\tau_i \leq 0{,}28$	$0{,}28 < \tau_i \leq 0{,}85$	$\tau_i > 0{,}85$
-	0,32	0,33	0,35
0	0,34	0,32	0,34
+	0,32	0,32	0,36

Tabell 3.4: Virkning av y_{i-1} på τ_i

4
ACM-ACD-modellen

Som vi så i kapittel 3 har de høyfrekvente dataene flere spesielle egenskaper. De viktigste er

i) diskrete prisendringer, se 3.1
ii) negativ lag(1)-korrelasjon, se 3.2
iii) sammenheng mellom tidsavstander og prisendringer, se 3.5
iv) klynger i justerte tidsavstander, se 3.4

Punkt iii) forteller oss at det er en sammenheng mellom tidsavstander og prisendringer, og dette er årsaken til at Russell og Engle (2005) foreslår å modellere prisendringene y_i og tidsavstandene τ_i simultant.

Vi ønsker en modell for simultanfordelingen til de diskrete prisene og tid mellom handler betinget på den bivariate filtrasjonen av prisendringer og tidsavstander $f(y_i, \tau_i | y_{1:i-1}, \tau_{1:i-1})$ der $y_{1:i-1} = (y_{i-1}, y_{i-2}, \ldots, y_1)$ og $\tau_{1:i-1} = (\tau_{i-1}, \tau_{i-2}, \ldots, \tau_1)$.

Vi dekomponerer simultantettheten for y_i, τ_i i to marginalfordelinger etter formelen

$$\mathbb{P}(A \cap B|C) = \mathbb{P}(A|B \cap C)\,\mathbb{P}(B|C)$$

og får

$$f(y_i, \tau_i | y_{1:i-1}, \tau_{1:i-1}) = g(y_i | y_{1:i-1}, \tau_{1:i}) q(\tau_i | y_{1:i-1}, \tau_{1:i-1}).$$

Der q(.) modelleres med Autoregressive Conditional Duration (ACD) modell, se 4.1, mens g(.) modelleres med Autoregressive Conditional Multinomial (ACM) modell, se 4.2.

4.1 ACD-modellen

Vi ønsker å modellere de sesongjusterte tidsavstandene τ_i. Den enkleste modellen får vi dersom vi antar at tidsavstandene er eksponentialfordelt med konstant forventning ψ, dvs. $\exp(\psi)$-fordelt, men dette er ikke i tråd med funnene våre fra 3.4. Ut fra tabell 3.2 på forrige side ser vi nemlig at $\psi_i = \mathbb{E}(\tau_i | \mathcal{F}_{i-1})$ er avhengig av τ_{i-1}. Vi ønsker derfor å

modellere

$$q(\tau_i|y_{1:i-1}, \tau_{1:i-1}) = \frac{1}{\psi_i} \exp\Big(- \frac{\tau_i}{\psi_i}\Big). \tag{4.1}$$

Etter at vi har fjernet sesongen i tidsavstandene, har vi altså fremdeles struktur i tidsavstandene. Strukturen er svært lik den man finner i volatilitet, der det er et kjent fenomen at man finner struktur. Høy volatilitet en dag betyr stor sannsynlighet for høy volatilitet også neste dag. Dette er omtalt i Tsay (2005, s. 103). Tilsvarende vil man i tidsavstandene i høyfrekvente data finne at stor tidsavstand mellom to handler betyr stor sannsynlighet for at tidsavstanden mellom de to neste handlene også er stor. Dette fører til at vi modellerer tidsavstandene med modeller som er svært lik modeller brukt i volatilitetsmodellering, og vi ser først på en GARCH-modell. Denne modellen er beskrevet nærmere i Tsay (2005, s. 113-122).

En GARCH-modell kan skrives på formen

$$a_i = \sigma_i \epsilon_i, \quad \sigma_i^2 = \omega + \sum_{j=1}^{m} \alpha_j a_{i-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{i-j}^2,$$

der $\{\epsilon_i\}$ iid. stokastiske variabler med foventning 0 og varians 1. Vi har også restriksjonene $\alpha_0 > 0,\ \alpha_j \geq 0\ (j = 1, \ldots, m),\ \beta_j \geq 0\ (j = 1, \ldots, q)$ og $\sum_{j=1}^{m \vee q} (\alpha_j + \beta_j) < 1$.

For å modellere tidsavstandene brukes ACD-modellen. Denne modellen er beskrevet i Tsay (2005, s. 227-236) samt Engle og Russell (1998) og defineres nedenfor. Modellen kalles autoregressive conditional duration (ACD) siden den betingede forventningen til tidsavstandene vil avhenge av tidligere tidsavstander.

Definisjon 4.1.1: ACD(m,q)-modell

$$\tau_i = \psi_i \varepsilon_i, \quad \psi_i = \omega + \sum_{j=1}^{m} \alpha_j \tau_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j}, \tag{4.2}$$

der $\{\epsilon_i\}$ er iid. stokastiske variabler med fordeling exp(1).

Vi har at $\psi_i = \mathbb{E}(\tau_i|\mathcal{F}_{i-1})$ der $\mathcal{F}_{i-1} = \sigma\{\tau_s, s < i\}$ er informasjonen tilgjengelig ved tidspunkt t_i. Parametre i modellen er ω, α_j $(j = 1, \ldots, m)$ og β_j $(j = 1, \ldots, q)$.

Parametrene må velges slik at $\psi_i \geq 0$. Velger vi å ha ikke-negativitetskrav på parametrene, slik vi har i GARCH-modellen ovenfor, vil kravet $\psi_i \geq 0$ være oppfylt. Russell og Engle (2005) vil også ha med effekten av y_{i-1} på τ_i i modellen. Drar vi inn forklaringsvariabelen y_i, som tar både positive og negative verdier, blir det vanskelig å finne krav som gjør at $\psi_i \geq 0$ alltid er oppfylt. Derfor lanseres Nelson-form ACD-modellen, der $\psi_i \geq 0$ alltid er oppfylt uavhengig av hvilke parameterverdier vi har.

Modellen defineres som følger:

Definisjon 4.1.2: Nelson-form ACD(u,v,w)-modell

$$\tau_i = \psi_i \epsilon_i, \quad \log(\psi_i) = \omega + \sum_{j=1}^{u} \alpha_j \varepsilon_{i-j} + \sum_{j=1}^{v} \beta_j \log(\psi_{i-j}) + \sum_{j=1}^{w} (\rho_j y_{i-j} + \zeta_j y_{i-j}^2), \quad (4.3)$$

der $\{\epsilon_i\}$ er iid. stokastiske variabler med fordeling exp(1).

Vi har at $\psi_i = \mathbb{E}(\tau_i | \mathcal{F}_{i-1})$ med $\mathcal{F}_{i-1} = \sigma\{\tau_s, y_s, s < i\}$ er informasjonen tilgjengelig ved tidspunktet t_i. Parametre i modellen er ω, α_j $(j = 1, \ldots, u)$, β_j $(j = 1, \ldots, v)$, ρ_j og ζ_j $(j = 1, \ldots, w)$.

En Nelson-form ACD-modell skiller seg fra ACD-modellen i (4.2) ved at:
- logaritmen av ψ_i inngår
- ε_i inngår på høyresiden
- forklaringsvariablene prisendring og prisendring kvadrert inngår

Vi kan også skrive modellen på formen

$$\psi_i = \exp\left(\omega + \sum_{j=1}^{u} \alpha_j \varepsilon_{i-j} + \sum_{j=1}^{v} \beta_j \log(\psi_{i-j}) + \sum_{j=1}^{w} (\rho_j y_{i-j} + \zeta_j y_{i-j}^2)\right). \quad (4.4)$$

Vi ser da at $\psi_i \geq 0$ alltid vil være oppfylt, og vi trenger ikke noen restriksjoner på parametrene for at ikke-negativitetskrav skal være oppfylt. Nelson-form sikrer oss altså at ikke-negativitetskravet er oppfylt selv om vi tar med ekstra forklaringsvariabler, i dette tilfellet tidligere prisendringer.

Ved simulering og estimering av modellen kommer vi frem til at y_{i-j} ikke passer inn i modellen ovenfor, se 4.5. Dette så vi også ved hjelp av krysstabell, se 3.5. Ved tilpasning til reelle data, se kapittel 5, ser vi derfor på modellen

$$\log(\psi_i) = \omega + \sum_{j=1}^{u} \alpha_j \varepsilon_{i-j} + \sum_{j=1}^{v} \beta_j \log(\psi_{i-j}).$$

Vi vil i fortsettelsen omtale dette som ACD(u, v)-modellen. Vi kunne alternativt ha benyttet ACD-modellen på formen (4.2), men ved å bruke Nelson-form slipper vi å tenke på parameterrestriksjoner.

4.1.1 Stasjonaritet Nelson-form ACD-modellen

Bauwens og Giot (2000) hevder at $|\beta_1 + \ldots + \beta_v| < 1$ er et nødvendig krav for kovariansstasjonaritet. I denne artikkelen nevnes det også at det ikke er funnet noe tilstrekkelig krav for kovariansstasjonaritet.

4.2 ACM-modellen

Som nevnt har vi diskrete prisendringer, se 3.1, og vi ønsker å modellere

$$g(y) = \begin{cases} \pi_{-1} & y < 0 \\ \pi_0 & y = 0 \\ \pi_1 & y > 0, \end{cases} \tag{4.5}$$

der y er prisendring. Ved å skrive fordelingen ovenfor på en annen form, vil vi se at det er en multinomisk fordeling. Vi definerer først en multinomisk fordeling, slik det er gjort i Dobson (2002, s. 135-136).

Definisjon 4.2.1: Multinomisk fordeling

Vi har at $x = (x_1, \ldots, x_J)'$ er mult(n, π), der $\pi = (\pi_1, \ldots, \pi_J)'$ dersom

$$f(x|n) = \frac{n!}{(x_1)! \ldots (x_J)!}(\pi_1)^{x_1} \ldots (\pi_J)^{x_J} \text{ med } \sum_{j=1}^{J} x_j = n \text{ og } \sum_{j=1}^{J} \pi_j = 1.$$

Vi definerer indikatorvektoren x:

$$x = \begin{cases} (1,0,0)' & y < 0 \\ (0,1,0)' & y = 0 \\ (0,0,1)' & y > 0. \end{cases} \tag{4.6}$$

Setter vi $n = 1$ i definisjonen ovenfor og kaller de tre tilstandene for $j = -1, 0, 1$ får vi at

$$f(x|1) = \frac{1}{(x_{-1})!(x_0)!(x_1)!}(\pi_{-1})^{x_{-1}}(\pi_0)^{x_0}(\pi_1)^{x_1}$$

$$= \begin{cases} \pi_{-1} & x = (1,0,0)' \Leftrightarrow y < 0 \\ \pi_0 & x = (0,1,0)' \Leftrightarrow y = 0 \\ \pi_1 & x = (0,0,1)' \Leftrightarrow y > 0, \end{cases}$$

og vi ser at fordelingen (4.5) kan skrives som en multinomisk fordeling.

Som nevnt i 3.2 har vi negativ lag(1)-korrelasjon i prisendringen y. Ut fra tabell 3.1 på side 16 ser vi at fordelingen til y_i vil avhenge av y_{i-1}. Vi ønsker derfor å modellere den betingede fordelingen:

$$g(y_i|y_{1:i-1}, \tau_{1:i}) = \begin{cases} \pi_{i,-1} & y_i < 0 \\ \pi_{i,0} & y_i = 0 \\ \pi_{i,1} & y_i > 0. \end{cases} \tag{4.7}$$

Vi starter med en første ordens ikke-parametrisk Markovanalyse:

$$(\pi_{i,-1}, \pi_{i,0}, \pi_{i,1})' = \pi_i = P' x_{i-1}.$$

Her er x_i den samme indikatorvektoren som i (4.6) og P er en overgangsmatrise som må tilfredsstille:

i) alle elementer er ikke-negative
ii) alle rader summerer seg til én

En modell på denne formen tar hensyn til at vi har diskrete prisendringer og negativ lag(1)-korrelasjon, men vi må også ta hensyn til at det er en sammenheng mellom tidsavstander og prisendringer, som vi så i 3.5. Mer generelt vil altså P variere med informasjonen tilgjengelig ved $i-1$, som f.eks. flere lag av x, tidligere verdier av π og tidligere tidsavstander τ. Vi ønsker å unngå de to restriksjonene på matrisen P, men samtidig ønsker vi at elementene i vektoren π_i skal være ikke-negative og summere seg opp til én. For å få dette til bruker vi en logistisk transformasjon. Dette bygger på ideene i logistisk regresjon, som er omtalt i tillegg A. I fortsettelsen definerer vi $\log(\pi)$ og $\exp(\pi)$ som logaritmefunksjonen og eksponentialfunksjonen brukt på hvert av elementene i vektoren π.

Vi modellerer

$$h(\pi_i) = P' x_{i-1} + c, \tag{4.8}$$

der

$$\begin{aligned} h(\pi_i) &= \log\left(\frac{\pi_i}{1-\iota'\pi_i}\right), \\ \iota' &= (1,1) \end{aligned} \tag{4.9}$$

og

$$x_i = \begin{cases} (1,0)' & y_i < 0 \\ (0,0)' & y_i = 0 \\ (0,1)' & y_i > 0. \end{cases}$$

Dette gir oss at

$$\begin{aligned} \log\left(\frac{\pi_i}{1-\iota'\pi_i}\right) &= P' x_{i-1} + c \\ \frac{\pi_i}{1-\iota'\pi_i} &= \exp(P' x_{i-1} + c). \end{aligned}$$

Løser vi for π_i får vi at

$$(\pi_{i,-1}, \pi_{i,1})' = \pi_i = \frac{\exp(P' x_{i-1} + c)}{1 + \iota' \exp(P' x_{i-1} + c)}, \tag{4.10}$$

der $\pi_{i,0} = 1 - (\pi_{i,-1} + \pi_{i,1})$.

Vi generaliserer nå (4.8) til å avhenge av mer informasjon, og dette gir oss ACM-modellen.

Definisjon 4.2.2: ACM(p,q)-modell

$$h(\pi_i) = c + \sum_{j=1}^{p} A_j(x_{i-j} - \pi_{i-j}) + \sum_{j=1}^{q} B_j h(\pi_{i-j}) + \chi \log((\tau_i, \tau_{i-1})'), \tag{4.11}$$

der

$$g(x_i|x_{1:i-1}, \tau_{1:i}) = \begin{cases} \pi_{i,-1} & x_i = (1,0)' \\ 1 - (\pi_{i,-1} + \pi_{i,1}) & x_i = (0,0)' \\ \pi_{i,1} & x_i = (0,1)'. \end{cases}$$

Her er $h(\pi_i)$ gitt ved (4.9), mens π_i er en vektor slik som i (4.10). Parametre i modellen er $c = (c_1, c_2)'$,

$$A_j = \begin{bmatrix} a_{11}^j & a_{12}^j \\ a_{21}^j & a_{22}^j \end{bmatrix} (j = 1, \ldots, p),\ B_j = \begin{bmatrix} b_{11}^j & 0 \\ 0 & b_{22}^j \end{bmatrix} (j = 1, \ldots, q) \text{ og } \chi = \begin{bmatrix} \chi_{11} & \chi_{12} \\ \chi_{21} & \chi_{22} \end{bmatrix}.$$

Vi merker oss at vi i (4.11) kan løse for π_i, ekvivalent til (4.10). Vi får da at

$$\pi_i = \frac{\exp\left(c + \sum_{j=1}^{p} A_j(x_{i-j} - \pi_{i-j}) + \sum_{j=1}^{q} B_j h(\pi_{i-j}) + \chi \log((\tau_i, \tau_{i-1})')\right)}{1 + \iota' \exp\left(c + \sum_{j=1}^{p} A_j(x_{i-j} - \pi_{i-j}) + \sum_{j=1}^{q} B_j h(\pi_{i-j}) + \chi \log((\tau_i, \tau_{i-1})')\right)}. \tag{4.12}$$

4.2.1 Symmetrirestriksjoner i ACM-modellen

I tabell 3.1 på side 16 observerte vi symmetri. Dersom vi antar at prisdynamikkene er symmetriske for prisbevegelser oppover og prisbevegelser nedover, får vi symmetrirestriksjonene $c_1 = c_2$, $a_{11} = a_{22}$, $a_{12} = a_{21}$ og $b_{11} = b_{22}$.

4.2.2 Antall tilstander i ACM-modellen

Vi har valgt å modellere ACM-modellen med tre tilstander, mens Russell og Engle (2005) har valgt en modell med fem tilstander. Ved å modellere en ACM$(3,3)$-modell med fem tilstander og symmetrirestriksjoner får man 40 parametre, og det viser seg å være vanskelig å utføre sannsynlighetsmaksimeringsestimering. Ved å redusere antall tilstander, reduserer vi også antall parametre i modellen. Tilpasser vi en ACM$(3,3)$-modell med tre tilstander og symmetrirestriksjoner, får vi redusert antall parametre fra 40 til 14.

4.2.3 Stasjonaritet i ACM-modellen

Vi antar at A_j $(j = 1, \ldots, p)$ og B_j $(j = 1, \ldots, q)$ har full rang, og at tidsavstandene $\{\tau_i\}$ er stasjonære. Dersom alle verdiene av z som tilfredsstiller $|I_{k-1} - B_1 z - \ldots - B_q z^q| = 0$ ligger utenfor enhetssirkelen, vil π_i være strengt positiv. Dette vises i Russell og Engle (2005), og dette er et nødvendig krav for stasjonaritet. Her er I_{k-1} identitetsmatrisen med dimensjon $k-1$, k antall tilstander i modellen og p og q er antall lag. I vårt tilfelle vil antall tilstander k være lik 3.

4.3 Estimering av ACM-ACD-modellen

For å estimere modellen utfører vi sannsynnsynlighetsmaksimeringsestimering, og vi regner derfor ut likelihoodfunksjonen $\mathcal{L}(\theta)$, der

$$\begin{aligned}
\mathcal{L}(\theta) &= f(y_{1:n}, \tau_{1:n}|\theta) \\
&= f(y_n, \tau_n|y_{1:n-1}, \tau_{1:n-1}, \theta) f(y_{n-1}, \tau_{n-1}|y_{1:n-2}, \tau_{1:n-2}, \theta) \ldots f(y_1, \tau_1|\theta) \\
&= \prod_{i=2}^{n} f(y_i, \tau_i|y_{1:i-1}, \tau_{1:i-1}, \theta) f(y_1, \tau_1|\theta) \\
&= \prod_{i=2}^{n} g(y_i|y_{1:i-1}, \tau_{1:i}, \theta_1) q(\tau_i|y_{1:i-1}, \tau_{1:i-1}, \theta_2) g(y_1|\tau_1, \theta_1) q(\tau_1|\theta_2).
\end{aligned}$$

Her er θ_1 parametre i ACM-modellen og θ_2 er parametre i ACD-modellen. Vi maksimerer derfor $\ell(\theta) = \log(\mathcal{L}(\theta))$, der

$$\begin{aligned}
\ell(\theta) &= \sum_{i=2}^{n} \log(g(y_i|y_{1:i-1}, \tau_{1:i}, \theta_1)) + \log(g(y_1|\tau_1, \theta_1)) \\
&+ \sum_{i=2}^{n} \log(q(\tau_i|y_{1:i-1}, \tau_{1:i-1}, \theta_2)) + \log(q(\tau_1|\theta_2)) \\
&= \ell_1(\theta_1) + \ell_2(\theta_2),
\end{aligned}$$

der $\ell_1(\theta_1)$ og $\ell_2(\theta_2)$ er gitt ved henholdsvis den første og den andre summen.

4.3.1 State-Space Markovegenskap i ACM-ACD-modellen

I simulering og estimering av modellen bruker vi State-Space Markovegenskapen. Vi har at

$$\mathbb{P}(\xi_i|\xi_{i-1}, \ldots, \xi_1) = \mathbb{P}(\xi_i|\xi_{i-1}),$$

der $\xi_i = (\tau_i, \ldots, \tau_{i-u+1}, \psi_i, \ldots, \psi_{i-(u\vee v)+1})'$ i ACD-modellen, mens i ACM-modellen er $\xi_i = (x_i, \ldots, x_{i-p+1}, \pi_i, \ldots, \pi_{i-(p\vee q)+1})'$. For å beregne ψ_i i (4.4) trenger vi u lag av τ_i og $u \vee v$ lag av ψ_i. For å beregne π_i i (4.12) trenger vi bare kjenne p lag av x_i og $p \vee q$ lag av π_i. Vi kan derfor regne ut π_i og ψ_i rekursivt. Vi har nå at

$$\mathbb{P}(\xi_{1:n}) = \mathbb{P}(\xi_n|\xi_{1:n-1})\,\mathbb{P}(\xi_{n-1}|\xi_{1:n-2}) \ldots \mathbb{P}(\xi_1) = \mathbb{P}(\xi_n|\xi_{n-1})\,\mathbb{P}(\xi_{n-1}|\xi_{n-2}) \ldots \mathbb{P}(\xi_1).$$

Dette er en viktig egenskap som vi får bruk for i den nummeriske algoritmen som maksimerer likelihoodfunksjonen, se 4.6.

4.3.2 Algoritme for estimering av ACM-ACD(3,3)-(2,2)-modell

Vi ser først på estimering av ACM-delen og deretter på estimering av ACD-delen.

ACM-del

1) Setter startverdier for π_1, π_2 og π_3.
2) For $i = 4 : n$, der n er antall observasjoner
   a) Beregner ut fra startverdier og data

$$h(\pi_i) = c + \sum_{j=1}^{3} A_j(x_{i-j} - \pi_{i-j}) + \sum_{j=1}^{3} B_j h(\pi_{i-j}) + \chi \log((\tau_i, \tau_{i-1})')$$

   fra (4.11). Her er x_{i-j}, $j = 1, 2, 3$, τ_i og τ_{i-1} observerte data.
   b) Regner ved hjelp av $h(\pi_i)$ fra a) ut

$$\pi_i = \frac{\exp(h(\pi_i))}{1 + \iota' \exp(h(\pi_i))}.$$

   c) Begynner på nytt i a) med oppdatert verdi av π_i fra b).
3) Vi har nå regnet ut $\pi_1,\ldots,\pi_n$ rekursivt, og vi kan regne ut loglikelihoodfunksjonen:

$$\ell_{1,i}(\theta_1) = \log(g(y_i|y_{1:i-1}, \tau_{1:i})) = \begin{cases} \log(\pi_{i,-1}) & y_i < 0 \\ \log(1 - (\pi_{i,-1} + \pi_{i,1})) & y_i = 0 \\ \log(\pi_{i,1}) & y_i > 0, \end{cases}$$

   se (4.7).
4) Maksimer

$$\ell_1(\theta_1) = \sum_{i=1}^{n} \ell_{1,i}(\theta_1)$$

   med hensyn på θ_1 ved hjelp av BHHH-algoritmen, se 4.6. Her er θ_1 en vektor med parametrene i ACM-delen.

ACD-del

1) Setter startverdier for ψ_1 og ψ_2.
2) For $i = 3 : n$ der n er antall observasjoner
   a) Beregner fra startverdier og data ut

$$\psi_i = \exp\left(\omega + \sum_{j=1}^{2} \alpha_j \varepsilon_{i-j} + \sum_{j=1}^{2} \beta_j \log(\psi_{i-j}) + \sum_{j=1}^{2} (\rho_j y_{i-j} + \zeta_j y_{i-j}^2)\right),$$

   se (4.4). Her er $\varepsilon_i = \dfrac{\tau_i}{\psi_i}$, og τ_{i-j} og y_{i-j}, $j = 1, 2$ er observerte data.
   b) Begynner på nytt med oppdatert verdi av ψ_i fra a).
3) Vi har nå regnet ut $\psi_1,\ldots,\psi_n$ rekursivt, og vi kan regne ut loglikelihoodfunksjonen

$$\ell_{2,i}(\theta_2) = \log(q(\tau_i|y_{1:i-1}, \tau_{1:i-1}, \theta_2)) = \log\left(\frac{1}{\psi_i} \exp(-\tau_i/\psi_i)\right),$$

se (4.1).

4) Maksimer

$$\ell_2(\theta_2) = \sum_{i=1}^{n} \ell_{2,i}(\theta_2)$$

med hensyn på θ_2 ved hjelp av BHHH-algoritmen, se 4.6. Her er θ_2 en vektor med parametrene i ACD-delen.

Vi kan også estimere modellen ved å maksimere ACM-del og ACD-del simultant, dvs. at vi maksimerer

$$\ell(\theta) = \sum_{i=1}^{n} \ell_{1,i}(\theta_1) + \sum_{i=1}^{n} \ell_{2,i}(\theta_2).$$

Siden det ikke er noen felles parametre i ACM-delen og ACD-delen vil løsningen på optimeringsproblemet være den samme, men den nummeriske algoritmen vil påvirkes av at vi maksimerer modellen simultant.

4.4 Simulering av ACM-ACD-modellen

4.4.1 Algoritme for simulering av ACM-ACD(3,3)-(2,2)-modell

1) Setter startverdier ψ_2, ψ_3, τ_2, τ_3, y_2, y_3, x_1, x_2, x_3, π_1, π_2 og π_3.
2) For $i = 4 : n$, der n er antall observasjoner
   a) Regner fra startverdiene ut

$$\psi_i = \exp\left(\omega + \sum_{j=1}^{2} \alpha_j \varepsilon_{i-j} + \sum_{j=1}^{2} \beta_j \log(\psi_{i-j}) + \sum_{j=1}^{2} (\rho_j y_{i-j} + \zeta_j y_{i-j}^2)\right),$$

   der $\varepsilon_i = \dfrac{\tau_i}{\psi_i}$ fra (4.4). Her spares ψ_i til neste simulering og brukes i b).

   b) Fra (4.1) har vi at $\tau_i | y_{1:i-1}, \tau_{1:i} \sim \exp(\psi_i)$. Simulerer $\tau_i \sim \exp(\psi_i)$, der τ_i spares til neste simulering og brukes i c).

   c) Regner fra startverdiene og simulert τ_i-verdi ut vektoren

$$\pi_i = \frac{\exp\left(c + \sum_{j=1}^{3} A_j(x_{i-j} - \pi_{i-j}) + \sum_{j=1}^{3} B_j h(\pi_{i-j}) + \chi \log((\tau_i, \tau_{i-1})')\right)}{1 + \iota' \exp\left(c + \sum_{j=1}^{3} A_j(x_{i-j} - \pi_{i-j}) + \sum_{j=1}^{3} B_j h(\pi_{i-j}) + \chi \log((\tau_i, \tau_{i-1})')\right)},$$

   der $h(\pi_i)$ er definert i (4.9). Her spares π_i til neste simulering og brukes i d).

   d) Fra (4.7) har vi at

$$g(y_i | y_{1:i-1}, \tau_{1:i}) = \begin{cases} \pi_{i,-1} & y_i < 0 \\ \pi_{i,0} = 1 - (\pi_{i,-1} + \pi_{i,1}) & y_i = 0 \\ \pi_{i,1} & y_i > 0. \end{cases}$$

   Simulerer $y_i \sim \text{mult}(1, (\pi_{i,-1}, \pi_{i,0}, \pi_{i,1})')$.

e) Videre bestemmes x_i ut fra y_i:

$$x_i = \begin{cases} (1,0)' & y_i = -1 \\ (0,0)' & y_i = 0 \\ (0,1)' & y_i = 1 \end{cases}$$

Her spares y_i og x_i til neste simulering.

f) Starter ny simulering med a) og oppdaterte verdier av ψ_i fra a), τ_i fra b), π_i fra c) og y_i og x_i fra e).

Punkt b) gir oss den simulerte tidsavstanden τ_i, mens punkt d) gir oss den simulerte prisendringen y_i

4.5 Simuleringseksperiment ACM-ACD-modellen

Tabellen under viser resultatene vi får fra ett forsøk dersom vi simulerer en ACM$(2,1)$-modell og deretter forsøker å estimere modellen med de simulerte dataene. Tallene under parameterverdiene er parametrenes standardfeil basert på den empiriske informasjons-matrisen $\mathcal{J}$, der vi benytter at

$$\widehat{\text{Var}}(\hat{\theta}) = \mathcal{J}^{-1},$$

se Dobson (2002, s. 74-75).

	θ	100 000	25 000	15 000	10 000	5000	2000	1000
c_1	-0,060	-0,056	-0,057	-0,067	-0,043	-0,077	-0,156	-0,047
	0,004	0,005	0,009	0,013	0,011	0,023	0,070	0,035
a_{11}^1	-0,406	-0,436	-0,437	-0,370	-0,406	-0,363	-0,397	-0,502
	0,018	0,019	0,038	0,049	0,058	0,083	0,142	0,194
a_{21}^1	0,911	0,922	0,896	0,912	0,844	0,894	1,126	0,935
	0,014	0,014	0,028	0,037	0,045	0,064	0,100	0,148
a_{11}^2	0,510	0,530	0,524	0,487	0,501	0,481	0,476	0,608
	0,018	0,019	0,038	0,049	0,058	0,082	0,137	0,191
a_{21}^2	-0,760	-0,782	-0,753	-0,751	-0,694	-0,739	-0,887	-0,816
	0,015	0,016	0,031	0,041	0,049	0,070	0,130	0,156
b_{11}^1	0,950	0,954	0,955	0,944	0,963	0,939	0,871	0,960
	0,011	0,004	0,007	0,011	0,010	0,019	0,060	0,028
χ_{11}	0,221	0,218	0,213	0,238	0,195	0,165	0,215	0,300
	0,007	0,007	0,014	0,017	0,022	0,030	0,049	0,074
χ_{12}	-0,225	-0,221	-0,217	-0,238	-0,197	-0,174	-0,237	-0,297
	0,007	0,007	0,014	0,017	0,022	0,030	0,048	0,074
χ_{21}	0,211	0,210	0,208	0,227	0,219	0,212	0,257	0,225
	0,007	0,007	0,014	0,018	0,022	0,030	0,050	0,071
χ_{22}	-0,216	-0,214	-0,213	-0,229	-0,220	-0,219	-0,274	-0,224
	0,007	0,007	0,014	0,017	0,022	0,030	0,049	0,071

Første kolonne viser parameterverdiene vi har simulert med. For at dette skal være fornuftige verdier, har vi funnet disse verdiene ved å først tilpasse modellen til reelle data. Kolonne nummer to viser simuleringseksperimentet med 100 000 data, og de resterende kolonnene viser simuleringseksperimentet med færre data. Vi ser at ved bruk av datasett på under 10 000 data, vil estimeringen gi resultater som ikke er nær de virkelige parameterverdiene.

Korrelasjonsmatrisen til parametrene kan vi finne ut fra informasjonsmatrisen. Vi har at

$$C = \widehat{\mathrm{Var}}(\hat{\theta}) = \mathcal{J}^{-1}.$$

Korrelasjonsmatrisen er da gitt som

$$\widehat{\mathrm{corr}}(\hat{\theta}) = D^{-1/2} C D^{-1/2},$$

der D er matrisen som har samme diagonalelementer som matrisen C og alle andre elementer lik 0.

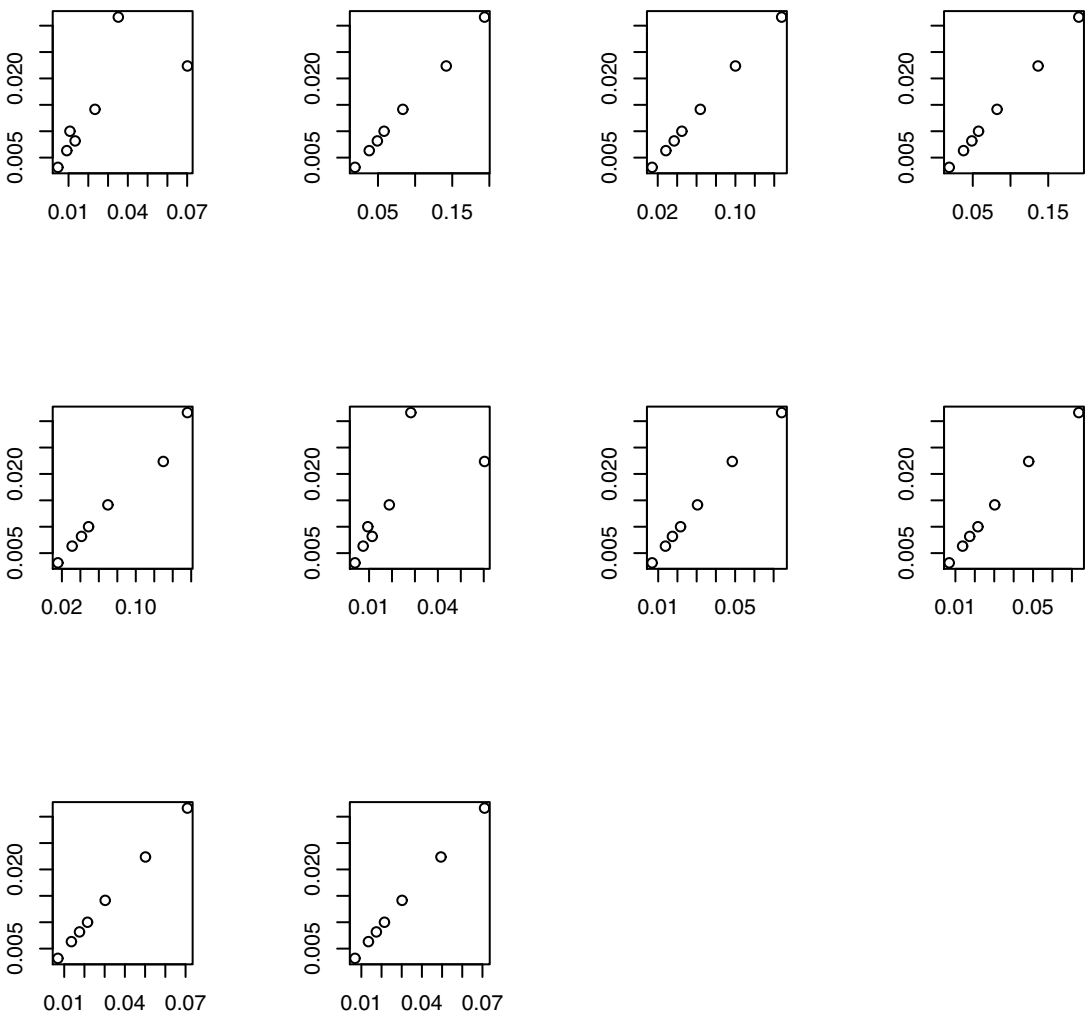
Korrelasjonsmatrisen nedenfor viser at det er sterk korrelasjon mellom parametrene c_1 og b_{11}, a^1_{11} og a^2_{11},a^1_{21} og a^2_{21}, χ_{11} og χ_{21} og mellom χ_{12} og χ_{22}. Korrelasjonsmatrisen er her hentet fra forsøket med 100 000 data.

	c_1	a^1_{11}	a^1_{21}	a^2_{11}	a^2_{21}	b_{11}	χ_{11}	χ_{21}	χ_{21}	χ_{22}
c_1	1	0,02	-0,02	-0,12	-0,39	**0,98**	0,02	0,00	0,03	0,00
a^1_{11}		1	0,46	**-0,95**	-0,37	0,01	0,02	-0,02	0,02	-0,02
a^1_{21}			1	-0,42	**-0,84**	-0,01	0,02	-0,02	0,02	-0,02
a^2_{11}				1	0,46	-0,11	-0,02	0,02	-0,02	0,01
a^2_{21}					1	-0,41	-0,02	0,02	-0,02	0,02
b_{11}						1	0,01	-0,02	0,02	-0,02
χ_{11}							1	**-0,99**	0,24	-0,23
χ_{12}								1	-0,23	0,24
χ_{21}									1	**-0,99**
χ_{22}										1

I Russell og Engle (2005) er det ikke beskrevet hvordan man har funnet standardfeilene til parametrene. I tabellen på forrige side har vi funnet estimatene fra informasjonsmatrisen. For å kunne avgjøre hvilke parametre som er signifikante, må vi forsikre oss om at vi har gode estimater av standardfeilene. For å teste dette kan vi sjekke om ulike teoretiske egenskaper holder. Vi skal ha at

$$\widehat{\mathrm{SD}}(\hat{\theta}) = n^{-1/2}\hat{\sigma}.$$

Ved å plotte $\widehat{\mathrm{SD}}(\hat{\theta})$ mot $n^{-1/2}$ skal vi da få en rett linje. Figuren på neste side viser standardfeil til hver av de 10 parametrene plottet mot $n^{-1/2}$. I følge teorien skal dette gi en rett linje, noe som viser seg å stemme godt, med unntak av enkelte av verdiene for parametrene c_1 og b_{11}. Vi ser også dette direkte ut fra tabellen ovenfor ved at standardfeilene ved 100 000 simuleringer er 10 ganger så små som ved 1000 simuleringer.



Figur 4.1: Standardfeil til hver av de 10 parametrene plottet mot $n^{-1/2}$. Øverste linje viser et slikt plot for c_1, a^1_{11}, a^1_{21} og a^2_{11}. Linjen i midten viser plot for a^2_{21}, b_{11}, χ_{11} og χ_{12}, mens nederste linje viser plot for χ_{21} og χ_{22}.

Vi kan også teste kvaliteten på standardfeilene ved å gjøre et gjentaksforsøk, og se hvor godt disse verdiene samsvarer med verdiene vi fikk ut fra informasjonsmatrisen. Vi gjør 100 simuleringer med 10 000 observasjoner i hver simulering og regner deretter ut estimatene ved å ta gjennomsnitt av parameterestimatene fra hver simulering, dvs. at vi regner ut

$$\frac{1}{m}\sum_{i=1}^{m}\hat{\theta}_i,$$

der m er antall gjentak dvs. 100. Tabellen til venstre på neste side viser resultatene vi får. Vi har også regnet ut standardfeilene og ser at disse stemmer godt overens med standardfeilene vi får ut fra informasjonsmatrisen. Dvs. at vi sammenligner

$$\frac{1}{m-1}\sum_{i=1}^{m}(\hat{\theta}_i - \bar{\hat{\theta}})^2$$

med $\widehat{\text{Var}}(\hat{\theta})$ fra informasjonsmatrisen. Ut fra dette eksperimentet kan det se ut som det er en tendens til at standardfeilene vi får ut ved gjentaksforsøket, er høyere enn verdiene vi får ut fra informasjonsmatrisen.

	θ	$\bar{\hat{\theta}}$
c_1	-0,060	-0,066
a_{11}^1	-0,406	-0,411
a_{21}^1	0,911	0,908
a_{11}^2	0,510	0,516
a_{21}^2	-0,760	-0,750
b_{11}^1	0,950	0,945
χ_{11}	0,221	0,222
χ_{12}	-0,225	-0,226
χ_{21}	0,211	0,216
χ_{22}	-0,216	-0,220

	Std.feil gjentak	Std.feil info.mat.
c_1	0,016	0,011
a_{11}^1	0,061	0,058
a_{21}^1	0,050	0,045
a_{11}^2	0,060	0,058
a_{21}^2	0,054	0,049
b_{11}^1	0,014	0,010
χ_{11}	0,022	0,022
χ_{12}	0,022	0,022
χ_{21}	0,022	0,022
χ_{22}	0,022	0,022

Tabellene nedenfor viser simuleringseksperiment med ACD-delen i modellen. Vi ser at ved bruk av datasett på under 5000 data vil estimeringen gi resultater som ikke er nær de virkelige parameterverdiene.

	θ	100 000	25 000	15 000	10 000	5000	2000	1000
ω	-0,0523	-0,0518	-0,0499	-0,0508	-0,0516	-0,0501	-0,0334	-0,0619
	0,0007	0,0012	0,0024	0,0029	0,0036	0,0056	0,0075	0,0125
α_1	0,0514	0,0509	0,0490	0,0497	0,0505	0,0488	0,0345	0,0615
	0,0007	0,0012	0,0023	0,0029	0,0035	0,0055	0,0078	0,0125
β_1	0,9888	0,9888	0,9897	0,9890	0,9894	0,9908	0,9924	0,9914
	0,0004	0,0007	0,0013	0,0017	0,0020	0,0029	0,0043	0,0050

Vi har også regnet ut korrelasjonsmatrisen, slik vi også gjorde for ACM-delen. Vi observerer en sterk negativ korrelasjon mellom konstantleddet ω og α_1 i den estimerte ACD-modellen.

	ω	α_1	β_1
ω	1	**-0,99**	0,42
α_1		1	-0,38
β_1			1

4.6 Nummeriske problemer ved estimering av modellen

For å estimere parametrene i modellen benytter vi BHHH-algoritmen. Dette er Newton-Raphson med hessematrisen approksimert med ytreproduktet av gradientvektoren, beskrevet i Berndt *et al.* (1974).

4.6.1 Newton-Raphson

Vi ønsker å maksimere loglikelihoodfunksjonene $\ell_1(\theta_1)$ og $\ell_2(\theta_2)$ der

$$\ell_1(\theta_1) = \sum_{i=1}^{n} \ell_{1,i}(\theta_1)$$

og

$$\ell_2(\theta_2) = \sum_{i=1}^{n} \ell_{2,i}(\theta_2).$$

Vi starter med å se på Newtonsmetode, slik den er definert i Gerald og Wheatley (2004, s. 42-43). Vi ønsker å finne x som oppfyller $f(x) = 0$. Vi har at

$$f'(x_0) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Setter vi $f(x_1) = 0$ og løser for x_1 får vi

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Dette gir oss Newtonsmetode

$$x_m = x_{m-1} - \frac{f(x_{m-1})}{f'(x_{m-1})},\ m = 1, 2, \ldots$$

For å maksimere en loglikelihoodfunksjon må vi løse ligningen

$$\frac{\partial \ell}{\partial \theta} = U(\theta) = 0.$$

Vi lar $x_m = \theta_m$ og $f(x_m) = U(\theta_m)$. Vi får da at

$$\theta_m = \theta_{m-1} - \frac{U(\theta_{m-1})}{U'(\theta_{m-1})}.$$

Vi ønsker nå å generalisere til tilfellet hvor f er k-dimensjonal. Vi har nå at

$$x_m = x_{m-1} - (\nabla f(x_{m-1}))^{-1} f(x_{m-1}),$$

der $x_m \in \mathbb{R}^k$.

Setter vi $x_m = \theta_m$ og $f(x_m) = \nabla \ell(\theta_m)$ får vi at

$$\theta_m = \theta_{m-1} - (\nabla^2 \ell(\theta_{m-1}))^{-1} \nabla \ell(\theta_{m-1}).$$

For å bruke denne iterasjonen direkte må vi beregne de andrederiverte av loglikelihoodfunksjonen. Ved å utnytte Markovegenskapen i modellen kan vi unngå dette. Hvis observasjonene våre kan oppfattes som data fra en homogen Markovmodell, har vi at

$$f(x, \theta) = f(x_n, \theta | x_{n-1}) \ldots f(x_2, \theta | x_1) f_0(x_1, \theta).$$

Vi ser i 4.3.1 at modellen har denne egenskapen. Observasjonene er fast og vi definerer

$$g_0(\theta) \stackrel{\text{def}}{=} \log(f_0(x_1, \theta))$$
$$g_j(\theta) \stackrel{\text{def}}{=} \log(f(x_j, \theta | x_{j-1})),\ j \geq 1.$$

Da får vi med n fast

$$\ell(\theta) = \sum_{j=0}^{n} g_j(\theta)$$
$$\nabla\ell(\theta) = \sum_{j=0}^{n} \nabla g_j(\theta).$$

Vi kan nå tilnærme $\nabla^2\ell(\theta)$ ved å bruke at

$$\nabla^2\ell(\theta) = \frac{1}{n}\sum_{j=0}^{n} \nabla g_j(\theta)(\nabla g_j(\theta))'.$$

4.6.2 Valg av startverdier

Dårlige initialverdier gjør at algoritmen bruker lenger tid på å finne en løsning. Bruker vi simulerte data, er det ikke noe problem å bestemme initialverdiene, men ved bruk av reelle data vil dette være et problem. I estimeringen med reelle data har vi derfor først estimert modellen med lavest mulig orden og satt initialverdiene ved denne estimeringen til små verdier. Vi fikk da noen estimerte parameterverdier og brukte disse verdiene som initialverdier ved estimering av modell med høyere orden. Parameterne som kommer i tillegg ved høyere orden, får initialverdi lik null. På denne måten øker man med ett lag for hver estimering.

Algoritmen gir lokale maksimum. For å sjekke at vi har funnet et globalt maksimum, må vi kjøre algoritmen med flere ulike startverdier. Vi kan også simulere den estimerte modellen og se om de simulerte dataene har alle egenskapene til de reelle dataene brukt i estimeringen, se 5.4.2. Vi får da samtidig svar på hvor godt modellen passer til dataene.

5 Tilpasning av ACM-ACD-modellen til reelle data

5.1 Bearbeiding av data

De reelle dataene brukt i oppgaven er hentet fra Netfonds.no, og vi har sett på to datasett, ett for StatoilHydro i perioden 6. mars-3. juni 2008 og ett datasett for DNO i perioden 6. mars-26. juni 2008. Det første datasettet inneholder 107 379 observasjoner, mens det andre datasettet har 67 386 observasjoner. Figuren nedenfor viser et eksempel på hvordan datafilene ser ut. I dataene finner vi opplysninger om tidspunktet for handelen, prisen og volumet. Videre får vi også opplyst om det har vært en automatisk handel og hvilken megler som representerte henholdsvis kjøper og selger. For oss vil kun de to første kolonnene være interessante. De to første kolonnene representerer tidspunktet t_i og prisen p_i ved handel nummer i, og gjør det mulig for oss å regne ut tidsavstandene $\tau_i = t_i - t_{i-1}$ og prisendringene $y_i = p_i - p_{i-1}$.

time	price	quantity	board	source	buyer	seller
20080403T124004	158.2	400	Auto	trade	ND	MLI
20080403T124005	158.2	150	Auto	trade	ND	MLI
20080403T124007	158.2	1100	Auto	trade	ND	MLI
20080403T124007	158.2	300	Auto	trade	NEO	MLI
20080403T124031	158.1	4400	Auto	trade	DDB	ASC
20080403T124031	158.1	800	Auto	trade	GLI	ASC
20080403T124031	158.1	19800	Auto	trade	CDV	ASC
20080403T124031	158.1	30200	Auto	trade	CDV	CA
20080403T124031	158.1	5000	Auto	trade	GLI	CA
20080403T124031	158.1	4700	Auto	trade	GLI	CA
20080403T124031	158.1	550	Auto	trade	GLI	CA
20080403T124031	158.1	1350	Auto	trade	GLI	CA
20080403T124038	158.	250	Auto	trade	CA	MSI
20080403T124046	158.	3600	Auto	trade	CA	GLS
20080403T124046	158.	2000	Auto	trade	ND	GLS

Figur 5.1: Utdrag fra data StatoilHydro 3. april 2008

5.1.1 Fjerning av data

Dersom man studerer dataene, vil man se at det svært ofte forekommer at flere handler er registrert med nøyaktig samme tidspunkt. For å forstå dette fenomenet må man studere måten aksjer handles på. Hvis man ser på ordreboken, se figur 2.1 på side 10, ser man at det er 14 314 aksjer som er til salgs til prisen 151,10. Velger man en ordrestørrelse på 14 314 med en pris på 151,10, får man kjøpt disse aksjene umiddelbart, men det er viktig å merke seg at ordreboken forteller oss at disse 14 314 aksjene er fordelt på seks ulike salgsordrer. Dvs. at det er seks ulike personer som til sammen vil selge disse aksjene, og i det vi utfører kjøpsordren, blir det utløst seks handler med nøyaktig samme tidspunkt.

Ut i fra figuren på forrige side ser man at det på tidspunktet 12.40.31 er registrert åtte handler. Vi løser dette problemet ved å se på alle handler utført på samme tidspunkt som én handel. Vi lar prisen ved handelen være lik prisen ved den siste handelen på tidspunktet.

I estimeringen av modellen har vi i tillegg utelatt data fra den første halvtimen (9.00-9.30) og den siste halvtimen (16.00-16.30) av handelsdagen. Årsaken til dette er at man i disse tidsrommene har en åpnings- og en sluttauksjon der man får mange handler på svært kort tid, og dataene i dette tidsrommet skiller seg derfor fra resten av dataene.

5.1.2 Fjerning av sesong i tidsavstander

Som nevnt i 3.3 må sesongen fjernes fra τ_i. Sesongen fremkommer på figur 3.6 på side 18, som viser gjennomsnittlig tidsavstand ulike tidspunkter på dagen. For å plotte denne kurven har vi samlet data for en halv måned, sortert alle disse dataene i ulike grupper etter når på dagen handlene fant sted og beregnet gjennomsnittlig tidsavstand i hver gruppe.

Vi ønsker å justere dataene våre for sesong, og vi utfører en kubisk splineinterpolasjon av kurven på figur 3.6 på side 18, som beskrevet i Gerald og Wheatley (2004, s. 171). Dette gir oss ligningen

$$g_j(t) = a_j(t - t_j)^3 + b_j(t - t_j)^2 + c_j(t - t_j) + d_j.$$

Her bestemmes a_j, b_j, c_j og d_j som beskrevet i Gerald og Wheatley (2004, s. 170-174), mens $t_1, \ldots, t_n$ er punktene på figur 3.6 på side 18 og $t \in (t_j, t_{j+1})$. Justert tidsavstand kan da regnes ut som $\dfrac{\tau}{g_j(t)}$, der τ er ujustert tidsavstand og t er tidspunktet på dagen målt i sekunder.

5.2 Valg av modellorden

For å bestemme orden på modellen kan vi bruke et informasjonskriterium. De mest kjente er Akaikes informasjonskriterium (AIC) og Bayesiansk informasjonskriterium (BIC), også kalt Schwarz informasjonskriterium. Begge disse informasjonskriteriene er beskrevet i Tsay (2005, s. 41-43). Vi har at

$$\begin{aligned} \text{BIC} &= \log(n)k - 2\log(\mathcal{L}), \\ \text{AIC} &= 2k - 2\log(\mathcal{L}), \end{aligned}$$

der n er antall observasjoner, k er antall parametre og $\mathcal{L}$ er likelihoodfunksjonen.

Optimal modellorden er ordenen som minimerer informasjonskriteriet. Vi ser at vi får en straff ved å legge til flere parametre, og denne er lik $\log(n)k$ for BIC og $2k$ for AIC. Ved store datasett er straffen langt mindre for AIC enn for BIC, og AIC har en tendens til å gi for høy orden. Nedenfor har vi derfor valgt å bruke BIC-kriteriet. Ut i fra BIC-kriteriet nedenfor ser det ut som vi bør velge en ACM-ACD$(2,2)$-$(2,2)$-modell fremfor ACM-ACD$(1,1)$-$(2,2)$ og ACM-ACD$(3,3)$-$(2,2)$.

Modell	StatoilHydro	DNO
ACM-ACD$(1,1)$-$(2,2)$	392092	240398
ACM-ACD$(2,2)$-$(2,2)$	**391312**	**239849**
ACM-ACD$(3,3)$-$(2,2)$	391332	239892

Tabell 5.1: Utregnede verdier av BIC-kriteriet

5.3 Estimeringsresultater

I kapittel 4 så vi gjennom simuleringseksperiment at modellene krevde store mengder data for å gi gode resultater. For å få store nok datasett med reelle data må vi derfor lime sammen data fra flere dager. Når vi estimerer modellen, setter vi lagene lik en startverdi når vi kommer til et punkt i tidsrekken som tilsvarer starten på en ny dag, ettersom det er rimelig å anta at de første handlene en dag har lite å gjøre med de siste handlene dagen før.

Tabellen nedenfor viser resultatene ved estimering av en ACM-ACD$(2,2)$-$(2,2)$-modell. Tallene i parentes er parametrenes standardfeil, og vi ser at alle parametre med unntak av b_{11}^2 er signifikante.

	StatoilHydro	DNO
c_1	-0,060 (0,004)	-0,087 (0,007)
a_{11}^1	-0,406 (0,018)	-0,040 (0,019)
a_{21}^1	0,911 (0,014)	0,623 (0,017)
a_{11}^2	0,510 (0,018)	0,213 (0,020)
a_{21}^2	-0,760 (0,015)	-0,443 (0,019)
b_{11}^1	0,950 (0,011)	0,890 (0,020)
b_{11}^2	**0,002** (0,011)	**0,004** (0,018)
χ_{11}	0,221 (0,007)	0,195 (0,008)
χ_{12}	-0,225 (0,007)	-0,196 (0,008)
χ_{21}	0,211 (0,007)	0,199 (0,008)
χ_{22}	-0,216 (0,007)	-0,202 (0,008)
ω	-0,033 (0,001)	-0,061 (0,002)
α_1	0,091 (0,002)	0,114 (0,002)
α_2	-0,059 (0,002)	-0,053 (0,003)
β_1	1,223 (0,020)	1,113 (0,024)
β_2	-0,228 (0,020)	-0,119 (0,024)

Tabell 5.2: Estimeringsresultater ACM-ACD$(2,2)$-$(2,2)$-modell

Ved estimering av ACM-ACD-modellen er ikke y_i og y_i^2 i (4.3) signifikante. Dersom vi forsøker å simulere en modell med y_i og y_i^2 inkludert, og deretter forsøker å estimere, vil de estimerte verdiene ligge lagt unna sine virkelige verdier, selv ved bruk av store datamengder. Dette tyder på at y_i og y_i^2 ikke passer så godt inn i modellen, og vi har derfor utelatt disse leddene i estimeringen ovenfor.

5.3.1 Stasjonaritet i modellene

Vi ønsker å sjekke om de estimerte modellene oppfyller de nødvendige kravene for stasjonaritet. Vi ser først på ACM-delen, og løser $|I_2 - B_1 z - B_2 z^2| = 0$ for hver av de to estimerte modellene, se 4.2.3. For StatoilHydro-modellen får vi røttene $z_1 = -476{,}05$ og $z_2 = 1{,}05$, mens vi for DNO-modellen finner at $z_1 = -223{,}62$ og $z_2 = 1{,}12$. Vi ser at alle røttene ligger utenfor enhetssirkelen, og det nødvendige kravet for stasjonaritet i ACM-delen er oppfylt.

Vi ser deretter på ACD-delen, og vi løser $|1 - \beta_1 z - \beta_2 z^2| = 0$. For StatoilHydro-modellen får vi røttene $z_1 = 4{,}36$ og $z_2 = 1{,}01$, mens vi for DNO-modellen får røttene $z_1 = 8{,}35$ og $z_2 = 1{,}01$. Vi ser at alle røttene ligger utenfor enhetssirkelen. Vi kontrollerer også det nødvendige kravet nevnt i 4.1.1, dvs. $|\beta_1 + \beta_2| < 1$. For StatoilHydro-modellen finner vi at $|\beta_1 + \beta_2| = 0{,}995 < 1$, mens vi for DNO-modellen får at $|\beta_1 + \beta_2| = 0{,}994 < 1$, og vi ser at kravet er oppfylt for begge modellene.

5.4 Diagnostiske tester

Diagnostisk test av modellen er viktig for å vite om modellen passer bra til dataene våre. En diagnostisk test gir oss også et svar på om vi har funnet det globale maksimumet ved sannsynlighetsmaksimeringsestimeringen. Vi studerer først en tilnærming til diagnostisk test foreslått av Russell og Engle (2005) samt Bauwens *et al.* (2008, kap. 8), før vi studerer en alternativ måte å teste kvaliteten på modellen.

5.4.1 Test for autokorrelasjon i residualer

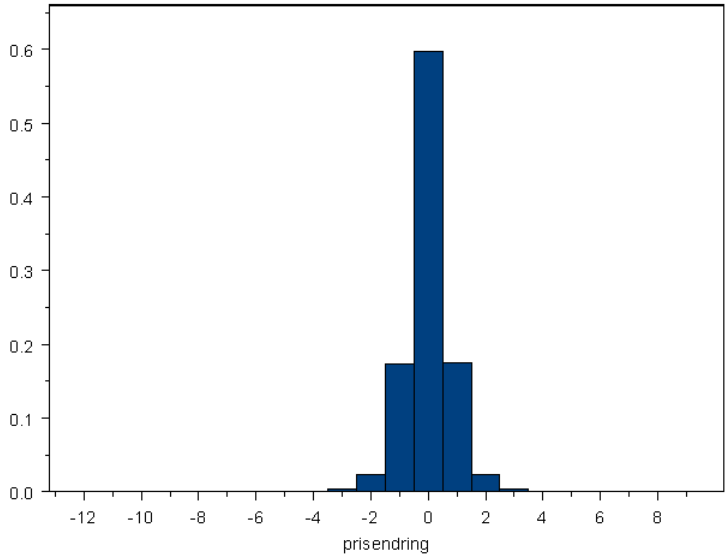
Russell og Engle (2005) beregner residualene i ACD-delen som $e_i = \frac{\tau_i}{\hat{\psi}_i}$. Disse residualene kan plottes og testes for autokorrelasjon. Problemet med denne tilnærmingen er at det er lite autokorrelasjon i dataene τ, derfor er det vanskelig å se om man får noe særlig mindre autokorrelasjon i residualene. Man kan teste for autokorrelasjon i residualene, men som følge av de store datamengdene vi bruker i modelleringen, vil vi lett forkaste en H_0-hypotese om ingen autokorrelasjon.

Residualene i ACM-delen beregner Russell og Engle (2005) ved å standardisere følgen av feil $v_i = x_i - \pi_i$. Man kan teste for autokorrelasjon i residualene ved å bruke en multivariat Portmanteau-observator, men vi får også her problemet med at vi lett forkaster på grunn av store datamengder. Siden vi her har multivariate residualer, er det ikke så lett å tolke plott av autokorrelasjon heller. Bauwens *et al.* (2008, kap. 8) løser problemet ved å regne ut testobservatoren fra rådataene og sammenligne dette med testobservatoren regnet ut fra residualene i den estimerte modellen. Blir testobservatoren redusert mye, har vi trolig en bra modell, men problemet er at vi ikke vet hvor mye testobservatoren må reduseres for at modellen skal være bra.

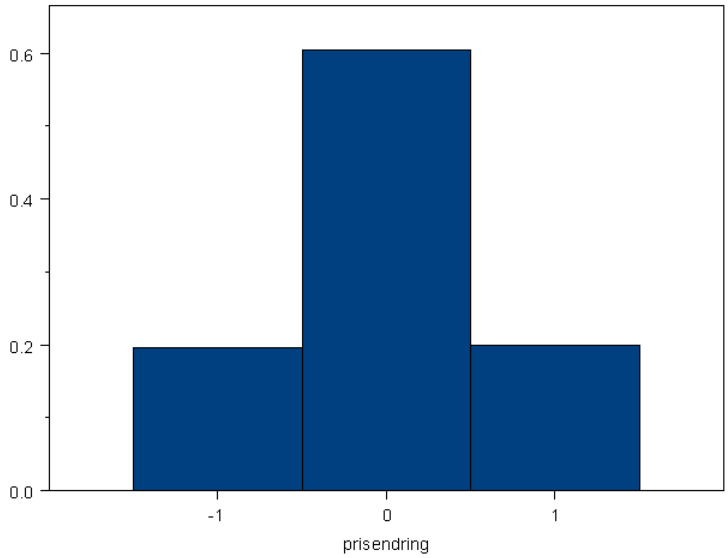
5.4.2 Simulering av estimert modell

For å sjekke modellen kan vi simulere den estimerte modellen og se om de simulerte dataene har alle egenskapene til de reelle dataene vi brukte i estimeringen. Vi har nedenfor gjort dette ved å simulere 100 000 observasjoner av den estimerte modellen for StatoilHydro-datasettet.

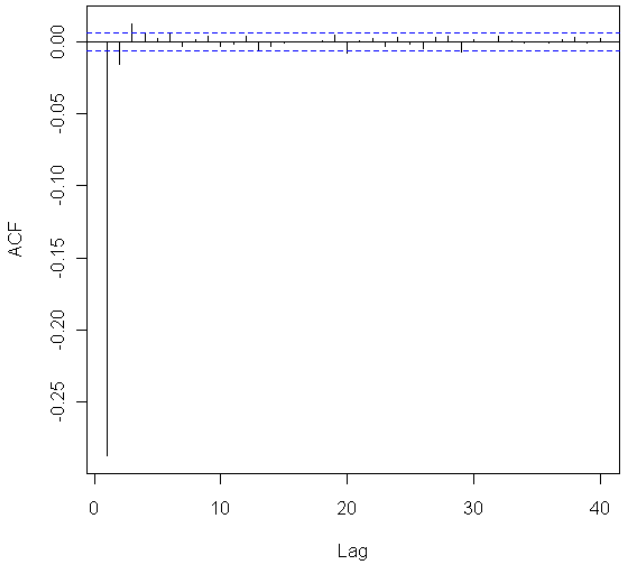
De to figurene under viser at den marginale fordelingen til prisendringene y er bevart i dataene. Vi ser også at autokorrelasjonen er lik i de simulerte og reelle dataene.



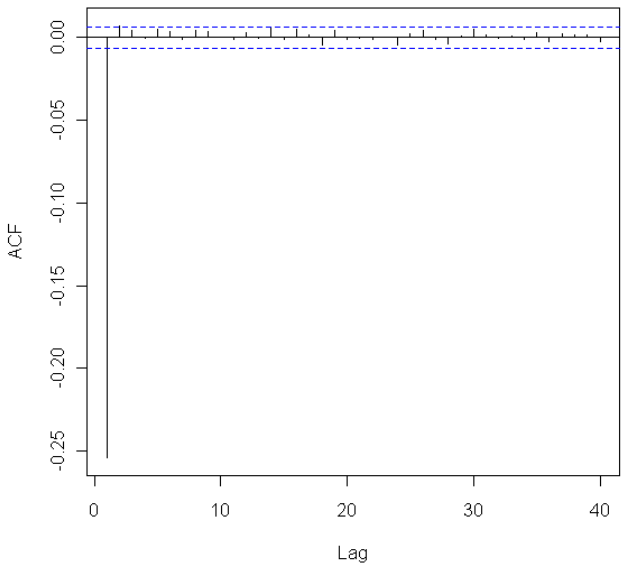
Figur 5.2: Empirisk fordeling til prisendringene i tickstørrelse reelle data



Figur 5.3: Empirisk fordeling til prisendringene i tickstørrelse simulerte data



Figur 5.4: Autokorrelasjon til prisendringene i reelle data



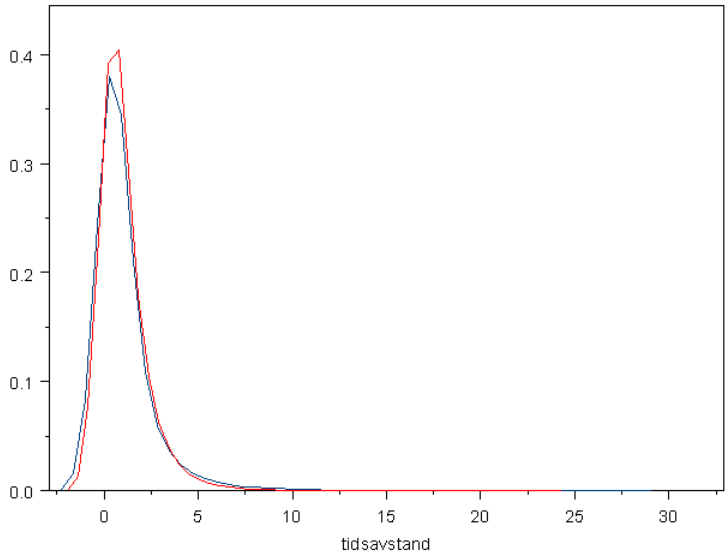
Figur 5.5: Autokorrelasjon til prisendringene i simulerte data

Tabellen viser virkningen av y_{i-1} på y_i i de simulerte dataene, mens tallene i parentes viser tilsvarende tall for reelle data. Vi ser at resultatene vi får for de reelle og simulerte dataene er svært like.

i-1\ i	-	0	+
−	0,10(0,11)	0,54(0,53)	0,36(0,36)
0	0,17(0,18)	0,65(0,64)	0,18(0,18)
+	0,36(0,36)	0,54(0,53)	0,10(0,11)

Nedenfor ser vi at virkningen av τ_i på y_i er bevart. Figuren viser at også den marginale fordelingen til tidsavstandene τ_i er bevart i de simulerte dataene.

	-	0	+
$\tau_i <= 0{,}28$	0,16(0,17)	0,68(0,66)	0,16(0,17)
$0{,}28 < \tau_i <= 0{,}85$	0,20(0,20)	0,60(0,60)	0,20(0,20)
$\tau_i > 0{,}85$	0,22(0,23)	0,56(0,54)	0,22(0,23)



Figur 5.6: Empirisk fordeling til tidsavstandene i sekunder reelle (blå kurve) og simulerte data(rød kurve)

5.5 Tolkning av modell og markedets mikrostruktur

Vi har i dette kapittelet tilpasset ACM-ACD-modellen til reelle data. Et relevant spørsmål er hvilken innsikt en slik modell gir oss. Her kommer temaet markedets mikrostruktur inn, som er beskrevet nærmere i O'Hara (1995). Studiet av høyfrekvente data gir oss en

dypere innsikt i hvordan handelen fungerer. En slik innsikt er nyttig både for dem som bestemmer hvordan reglene på børsen skal utformes, men også for den som handler i aksjer. Vi ser i dette avsnittet på et konkret eksempel.

5.5.1 Forbud mot kort salg

Høsten 2008 ble det forbud mot kort salg i enkelte aksjer på Oslo Børs. Med kort salg menes det at man har muligheten til å ha en negativ posisjon i en aksje. I praksis fungerer dette ved at man låner aksjen og selger den umiddelbart. Når eieren skal ha igjen aksjen på et senere tidspunkt, må investoren kjøpe aksjen på børsen før den leveres tilbake. Dette er beskrevet mer detaljert i Sandvik (2003, s. 129).

I følge O'Hara (1995, s. 168-170) vil man få en asymmetri i modellen dersom det eksisterer forbud mot kort salg. Dersom en investor mottar positiv informasjon om en aksje, vil han kjøpe aksjen umiddelbart. Dersom informasjonen er negativ, vil han benytte kort salg dersom det er mulig. Hvis kort salg ikke er mulig, vil det være umulig for en investor å benytte negativ informasjon. Investoren kan ved negativ informasjon ikke gjøre noe, og det vil gå lenger tid til neste handel enn ved positiv informasjon. Lengre tidsavstander er dermed et signal om fallende pris.

Den estimerte modellen for StatoilHydro-datsettet i 5.3 gir oss sammen med (4.11) at

$$
\begin{aligned}
h\begin{bmatrix}\pi_{i,-1}\\ \pi_{i,1}\end{bmatrix} &= \begin{bmatrix}-0{,}060\\ -0{,}060\end{bmatrix} + \begin{bmatrix}-0{,}406 & 0{,}911\\ 0{,}911 & -0{,}406\end{bmatrix}\begin{bmatrix}x_{i-1,-1}-\pi_{i-1,-1}\\ x_{i-1,1}-\pi_{i-1,1}\end{bmatrix}\\
&+ \begin{bmatrix}0{,}510 & -0{,}760\\ -0{,}760 & 0{,}510\end{bmatrix}\begin{bmatrix}x_{i-2,-1}-\pi_{i-2,-1}\\ x_{i-2,1}-\pi_{i-2,1}\end{bmatrix} + \begin{bmatrix}0{,}950 & 0\\ 0 & 0{,}950\end{bmatrix} h\begin{bmatrix}\pi_{i-1,-1}\\ \pi_{i-1,1}\end{bmatrix}\\
&+ \begin{bmatrix}0{,}221 & -0{,}225\\ 0{,}211 & -0{,}216\end{bmatrix}\log\begin{bmatrix}\tau_i\\ \tau_{i-1}\end{bmatrix}.
\end{aligned}
$$

Vi ønsker å studere virkningen av τ_i, og ser derfor på det siste leddet, og studerer derfor matrisen

$$\begin{bmatrix}0{,}221 & -0{,}225\\ 0{,}211 & -0{,}216\end{bmatrix}.$$

Den øverste raden viser virkningen av τ_i og τ_{i-1} på $\pi_{i,-1}$, mens den nederste raden viser virkningen av τ_i og τ_{i-1} på $\pi_{i,1}$. Her er $\pi_{i,-1}$ og $\pi_{i,1}$ de betingede sannsynlighetene for henholdsvis negativ og positiv prisendring, se (4.7). Dersom vi har asymmetri, vil første og andre rad være ulike.

Vi kan i dette tilfellet ut fra de estimerte verdiene ikke se noe tegn på asymmetri, noe som er rimelig ettersom kort salg var tillatt i StatoilHydro-aksjen i perioden vi har data for. Stemmer teorien, vil vi oppdage asymmetri ved tilpasning til data for en aksje hvor kort salg ikke er tillatt. Mer formelt kan man teste for symmetri ved å bruke en sannsynlighetskvotetest. Man må da estimere modellen med og uten symmetrirestriksjoner. Dette er beskrevet nærmere i Dobson (2002, s. 76-80).

5.6 Videre studier av modellen

5.6.1 Korrelasjon i parametre i modellen

I 4.5 observerte vi at enkelte av parametrene i modellen er sterkt korrelerte. Vi har i denne oppgaven ikke lykkes med å forklare hvorfor man finner denne korrelasjonen og om det har noen konsekvenser.

5.6.2 Sammenheng mellom volum og prisendring

Vi har i denne oppgaven begrenset oss til variablene prisendring og tidsavstand. Det kan også være interessant å se på sammenhengen mellom variablene volum og prisendring.

Investorer med informasjon vil forsøke å handle store volumer per transaksjon for å få mest mulig ut av informasjonen de sitter på. Disse store volumene sees på av de andre investorene som tegn på ny informasjon. Dermed kan man forvente større prisendringer etter store volum, enn etter små volum. Tolkningen er beskrevet i Bauwens *et al.* (2008, s. 186).

Det foreslås i Russell og Engle (2005) å inkludere volum på samme måte som tidsavstand i modellen. Modellen (4.11) blir da på formen

$$h(\pi_i) = c + \sum_{j=1}^{p} A_j(x_{i-j} - \pi_{i-j}) + \sum_{j=1}^{q} B_j h(\pi_{i-j}) + \chi \log((v_i, v_{i-1})'),$$

der v_i er volumet ved handel i.

5.6.3 Forbud mot kort salg

Som nevnt i 5.5 vil det være interessant å tilpasse ACM-ACD-modellen på et datasett hvor man har en aksje der kort salg ikke er tillatt. Vi har i denne oppgaven ikke hatt tilgang til et slikt datasett.

6
Alternativ modellering av aksjepris

Som også nevnt i begynnelsen av kapittel 4 har de høyfrekvente dataene flere spesielle egenskaper. De viktigste er

i) diskrete prisendringer, se 3.1
ii) negativ lag(1)-korrelasjon, se 3.2
iii) sammenheng mellom tidsavstander og prisendringer, se 3.5
iv) klynger i justerte tidsavstander, se 3.4

I de to foregående kapitlene studerte vi ACM-ACD-modellen. I dette kapittelet forsøker vi å lage en alternativ modell som kombinerer ideer fra Ait-Sahalia *et al.* (2005) og ACD-modellen til Russell og Engle (2005) studert i 4.1.

6.1 Brownsk bevegelse med støy

Vi starter med å definere en brownsk bevegelse slik det er gjort i Taylor og Karlin (1998, s. 476-477).

Definisjon 6.1.1: Brownsk bevegelse

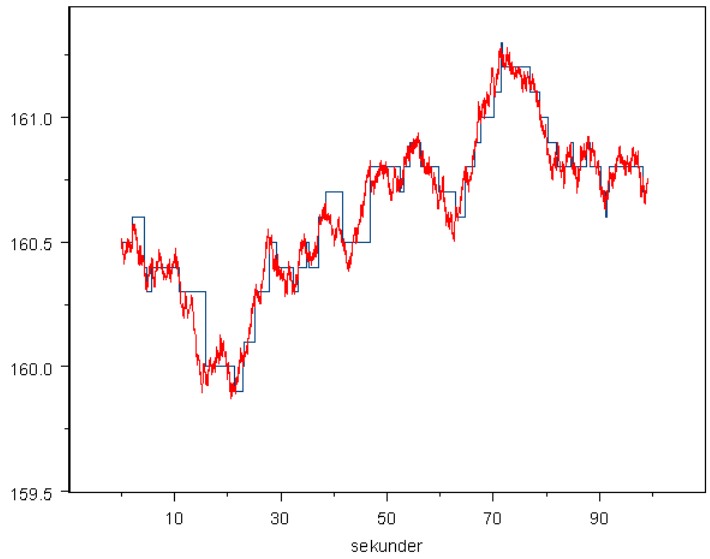
En brownsk bevegelse med diffusjonskoeffisient σ^2 er en stokastisk prosess $\{B(t);\ t \geq 0\}$ med de følgende egenskapene:

i) $B(s+t) - B(s) \sim \mathcal{N}(0, \sigma^2 t);\ \sigma^2 > 0$
ii) For alle par av disjunkte tidsintervall $(t_1, t_2],\ (t_3, t_4]$ med $0 \leq t_1 < t_2 \leq t_3 < t_4$ er $B(t_4) - B(t_3)$ og $B(t_2) - B(t_1)$ uavhengige stokastiske variabler, og tilsvarende for n disjunkte tidsintervall, hvor n er et vilkårlig positivt heltall.
iii) $B(0) = 0$, og $B(t)$ er en kontinuerlig funksjon av t.

Vi tenker oss at logaritmen til den virkelige prisen p_i^*, som vi ikke kan observere, følger en brownsk bevegelse. Dette er foreslått av bl.a. Ait-Sahalia *et al.* (2005). Videre

tenker vi oss at den virkelige prisen og tidsavstandene mellom handlene er uavhengige, og vi modellerer tidsavstandene med ACD-modellen, se 4.1.

Den brownske bevegelsen er en kontinuerlig funksjon av t, men vi observerer kun handler på diskrete tidspunkter. Prisen p_i vi observerer på disse tidspunktene vil være en avrunding til nærmeste pris i hele tickstørrelser, der tickstørrelse er den minste prisendring som kan tastes inn i handelssystemet, se definisjon 2.1.1 på side 11. Figuren under viser simulering av en brownsk bevegelse og den virkelige prisen som vi observerer.



Figur 6.1: Rød kurve viser virkelig pris, mens blå kurve viser observert pris

Ved å avrunde prisen til nærmeste tickstørrelse tar vi hensyn til punkt i) diskrete prisendringer, og ved å modellere tidsavstandene med ACD-modellen tar vi hensyn til punkt iv) klynger i tidsavstander. Det er ikke like opplagt hvordan punkt ii) og iii) tas hensyn til i modellen, og vi ser derfor nærmere på dette.

6.1.1 Autokorrelasjon i prisendringer

Simulerer vi modellen som beskrevet ovenfor, og plotter autokorrelasjonen til de observerte prisendringene, vil vi få et plot som ligner på det vi får på figur 5.4 på side 40. Årsaken til dette er at vi observerer en avrunding av den virkelige prisen. Vi observerer derfor en pris som er litt for høy eller litt for lav. Dette funnet gjør at vi kan diskutere om bidask-effekten beskrevet i 3.2 eksisterer, eller om den negative lag(1)-korrelasjonen kan forklares ved at vi observerer avrundinger av en kontinuerlig variabel.

6.1.2 Sammenheng mellom tidsavstander og prisendringer

Vi observerer prisen på diskrete tidspunkter og variansen til en brownsk bevegelse er gitt som $\sigma^2\tau$, se definisjon 6.1.1 på forrige side. Dvs. at variansen til den virkelige prisen øker med tidsavstanden. Dette fører til at større tidsavstand gir større sannsynlighet for observert prisendring.

6.1.3 Algoritme for simulering av modell

1) Setter startverdi for p_0^*.
2) For hver $i = 1, \ldots, n$, der n er antall observasjoner
   a) Simulerer tidsavstanden τ_i ved hjelp av ACD-modellen som beskrevet i 4.4.
   b) Simulerer $z \sim \mathcal{N}(0, 1)$.
   c) Regner ut
   $$\log(p_i^*) = \log(p_{i-1}^*) + \sigma\sqrt{\tau_i}z,$$
   der τ_i er gitt ved a) og z er gitt ved b). Fremgangsmåten for å simulere en brownsk bevegelse er nærmere beskrevet i Rasmus (2007, s. 168).
   d) Regner ut $p_i^* = \exp(\log(p_i^*))$.
   e) Regner ut observert pris p_i ved å runde av p_i^* til nærmeste pris i hele tickstørrelser.

6.1.4 Bestemmelse av parameter i modellen

I algoritmen ovenfor må parameteren σ velges. Denne parameteren kan vi bestemme ut fra reelle data. I modellen har vi at

$$y_i^* = \log(p_i^*) - \log(p_{i-1}^*)$$

og

$$\mathrm{Var}(y_i^*) = \sigma^2\tau_i.$$

Vi kan estimere σ^2 ut fra reelle data ved at

$$\hat{\sigma}^2 = \frac{1}{\bar{\tau}}\widehat{\mathrm{Var}}(y^*)$$

med

$$\widehat{\mathrm{Var}}(y^*) = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1},$$

der $y_i = \log(p_i) - \log(p_{i-1})$ er observasjoner.

7
Kointegrasjon og triangelarbitrasje

I de foregående kapitlene har vi sett på hvordan vi ved å modellere høyfrekvente data kan studere markedets mikrostruktur. Et spørsmål som ofte stilles i forbindelse med høyfrekvente data, er om dataene kan brukes til å tjene penger med. Her kommer temaet triangelarbitrasje inn, som er et fenomen man finner i valutamarkedet og studeres i dette kapittelet i tilknytning til kointegrasjonsmodellering. Vi ser derfor først på temaet kointegrasjon.

7.1 Kointegrasjon

7.1.1 Bivariate tilfellet

Vi starter med å definere enhetsrot slik det er gjort i Tsay (2005, s. 380-381):

Definisjon 7.1.1: Enhetsrot

Vi har en prosess $x_t = a_1 x_{t-1} + \ldots + a_p x_{t-p} + \epsilon_t$. Dersom prosessen har enhetsrot, dvs. at prosessen er $I(1)$, vil én av løsningene på $|1 - a_1 z - a_2 z^2 - \ldots - a_p z^p| = 0$ være $z = 1$, og de andre løsningene vil ligge utenfor enhetssirkelen.

Dette medfører at $(1-z)x_t$ er en stasjonær prosess, dvs. at den differensierte prosessen $x_t - x_{t-1}$ er $I(0)$.

Det er viktig å merke seg at en prosess som er $I(1)$, ikke trenger å være en tilfeldig gang. En tilfeldig gang kan skrives på formen

$$x_t = x_{t-1} + \epsilon_t,$$

der $\{\epsilon_t\}$ er en følge av iid. stokastiske variabler med $\mathbb{E}(\epsilon_t) = 0$. Vi ser på et eksempel på en prosess som har enhetsrot, men som ikke er en tilfeldig gang. Dette eksempelet er beskrevet i Pfaff (2008, s. 55).

Eksempel 7.1.2

Vi har at

$$x_t = x_{t-1} + \epsilon_t,\ \epsilon_t = \rho\epsilon_{t-1} + \xi_t,$$

hvor $|\rho| < 1$ og $\{\xi_t\}$ er en følge av iid. stokastiske variabler med $\mathbb{E}(\xi_t) = 0$. Siden $\{\epsilon_t\}$ ikke er en uavhengig følge, har vi ikke en tilfeldig gang, men vi ser at den differensierte prosessen

$$x_t - x_{t-1} = \epsilon_t$$

er en stasjonær prosess. Vi har dermed at x_t en $I(1)$-prosess.

Videre definerer vi kointegrasjon i det bivariate tilfellet, slik det er definert i Sørensen (2005):

Definisjon 7.1.3: Kointegrasjon for to skalare tidsrekker

$x_{1,t}$ og $x_{2,t}$ er kointegrerte hvis det eksisterer en parameter α slik at $u_t = x_{2,t} - \alpha x_{1,t}$ er en stasjonær prosess, der $x_{1,t}$ og $x_{2,t}$ begge er $I(1)$.

7.1.2 Generelle tilfellet

For det generelle tilfellet med flere enn to tidsrekker holder vi oss til tilnærmingen beskrevet i Tsay (2005, s. 380-381) og Pfaff (2008, s. 78-82). Vi definerer først en vektor-autoregressiv(VAR)-modell.

Definisjon 7.1.4: VAR(p)-modell med dimensjon k

$$x_t = \mu_t + \Phi_1 x_{t-1} + \ldots + \Phi_p x_{t-p} + \epsilon_t, \tag{7.1}$$

der x_t, μ_t og ϵ_t er $k \times 1$-vektorer, $\Phi_1, \ldots, \Phi_p$ er $k \times k$-matriser og $\{\epsilon_t\}$ er en følge av iid. stokastiske variabler med $\epsilon_t \sim \mathcal{N}(0, \Sigma)$.

Vi ser videre på enhetsrot i det generelle tilfellet.

Definisjon 7.1.5: Enhetsrot

En prosess har enhetsrot, dvs. at prosessen er $I(1)$, hvis $|\Phi(1)| = 0$ der $\Phi(B) = I - \Phi_1 B - \ldots - \Phi_p B^p$ og $|\Phi(z)| \neq 0$ for $|z| < 1$, dvs. at alle andre løsninger ligger utenfor enhetssirkelen.

For $p = 2$ kan (7.1) skrives som

$$x_t = \mu_t + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \epsilon_t.$$

Trekker fra x_{t-1} på begge sider og får da at:

$$\begin{aligned}
x_t - x_{t-1} &= \mu_t + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \epsilon_t - x_{t-1} \\
\Delta x_t &= \mu_t - (I - \Phi_1 - \Phi_2) x_{t-1} - \Phi_2 (x_{t-1} - x_{t-2}) + \epsilon_t \\
\Delta x_t &= \mu_t - \Phi(1) x_{t-1} - \Phi_2 \Delta x_{t-1} + \epsilon_t \\
\Delta x_t &= \mu_t + \Pi x_{t-1} + \Phi_1^* \Delta x_{t-1} + \epsilon_t,
\end{aligned}$$

der $\Pi = -\Phi(1)$ og $\Phi_1^* = -\Phi_2$.

Mer generelt får vi:

Definisjon 7.1.6: Kointegrasjonsmodell - Error correction model (ECM)

$$\Delta x_t = \mu_t + \Pi x_{t-1} + \Phi_1^* \Delta x_{t-1} + \ldots + \Phi_{p-1}^* \Delta x_{t-p+1} + \epsilon_t, \tag{7.2}$$

der $\Delta x_t = x_t - x_{t-1}$, $\Phi_j^* = -\sum_{i=j+1}^{p} \Phi_i$, $j = 1, \ldots, p-1$ og $\Pi = \alpha\beta' = -\Phi(1)$.

Ettersom de individuelle komponentene i x_t er $I(1)$-variabler, er venstre side i (7.2) stasjonær. Vi må derfor også ha at høyreside er stasjonær, og vi må ha at leddet Πx_{t-1} er stasjonært. Spørsmålet er hvilke betingelser vi må ha på matrisen Π slik at høyresiden er stasjonær. Som definert ovenfor er en prosess $I(1)$ hvis $|\Phi(1)| = 0$, det vil si at $\Pi = -\Phi(1)$ er singulær. For å avgjøre om vi har kointegrasjon kan vi derfor se på rangen til $k \times k$-matrisen Π. Vi ser på tre tilfeller:

1) $\text{Rang}(\Pi) = 0$ vil si at $\Pi = 0$ og at x_t ikke er kointegrert. Modellen reduseres til

$$\Delta x_t = \mu_t + \Phi_1^* \Delta x_{t-1} + \ldots + \Phi_{p-1}^* \Delta x_{t-p+1} + \epsilon_t.$$

2) $\text{Rang}(\Pi) = k$ vil si at $|\Phi(1)| \neq 0$ og x_t inneholder ikke noen enhetsrot. For at høyresiden skal være stasjonær må vi ha at x_t er stasjonær, og vi kan da studere x_t direkte.
3) $0 < \text{Rang}(\Pi) = m < k$ er det mest interessante tilfellet. Man kan da skrive Π som

$$\Pi = \alpha\beta',$$

hvor α og β er $k \times m$-matriser med $\text{rang}(\alpha) = \text{rang}(\beta) = m$, der $\alpha\beta' x_{t-1}$ er stasjonær. Vi får da en modell på formen

$$\Delta x_t = \mu_t + \alpha\beta' x_{t-1} + \Phi_1^* \Delta x_{t-1} + \ldots + \Phi_{p-1}^* \Delta x_{t-p+1} + \epsilon_t.$$

De m lineært uavhengige kolonnene til β er kointegrasjonsvektorene. Nullrommet til Π har da dimensjon lik $k - m$, og vi har dermed $k - m$ enhetsrøtter.

7.1.3 Kointegrasjon som økonomisk fenomen

Felles skjult faktor

Mange tidsrekker i finans og makroøkonomi er ikke stasjonære, men blir stasjonære dersom de blir differensiert. Dersom differansen mellom to slike skalare tidsrekker er stasjonær, sier vi at de er kointegrerte. Variabler som er kointegrerte, har en rekke interessante egenskaper. De må bl.a. ha en felles skjult faktor. Vi ser på et eksempel på to variabler som inneholder en felles skjult faktor.

Eksempel 7.1.7

Vi ser på et eksempel som er beskrevet i Pfaff (2008, s. 78). Vi lar

$$x_{1,t} = x_{1,t-1} + \epsilon_{1,t}$$

og

$$x_2 = 0{,}6x_1 + \epsilon_2,$$

hvor $\{\epsilon_{1,t}\}$ og $\{\epsilon_{2,t}\}$ er følger av iid. stokastiske variabler med $\mathbb{E}(\epsilon_{1,t}) = \mathbb{E}(\epsilon_{2,t}) = 0$. Dette gir kointegrasjonsvektoren $(1, -0{,}6)$' ettersom

$$x_2 - 0{,}6x_1 = \epsilon_2 \sim I(0)$$

er en stasjonær prosess.

Spuriøs regresjon

I makroøkonomi vil man ofte bruke modeller for å se på langsiktige sammenhenger. I de siste årene har error correction modeller(ECM), se definisjon 7.1.6 på forrige side, blitt populære for å se på slike sammenhenger, hvor kointegrasjon inngår. Anvendelsene som nevnes i Granger (2003) er på variabler som skatter, konsum, renter og statens utgifter. Error correction modeller(ECM) løser mange av vanskelighetene med spuriøs regresjon. Med spuriøs regresjon mener man at dersom man utfører en regresjon mellom to integrerte variabler vil man finne en sammenheng, selv om det ikke eksisterer noen sammenheng. Vi ser på et konkret eksempel på spuriøs regresjon.

Eksempel 7.1.8

Vi følger eksempelet som er beskrevet i Zivot og Wang (2006, s. 432-435). Vi ser på to uavhengige $I(1)$-prosesser som ikke er kointegrerte, $x_{1,t}$ og $x_{2,t}$ slik at

$$x_{i,t} = x_{i,t-1} + \epsilon_{i,t}, \text{ hvor } \{\epsilon_{i,t}\} \text{ er iid. og } \epsilon_{i,t} \sim \mathcal{N}(0,1),\ i = 1, 2.$$

Simulerer vi de to prosessene og utfører en lineær regresjon av x_1 på x_2 vil vi finne at de to variablene er korrelerte, selv om det ikke eksisterer noen sammenheng. Vi vil også se at det er sterk korrelasjon i residualene i regresjonen. En slik modell egner seg altså ikke for å modellere variabler som er $I(1)$. Dette er noe av grunnen til at man ønsker å studere kointegrasjonsmodeller.

Grangerkausalitet

Variabler som er kointegrerte, kan beskrives med en error correction modell(ECM), se definisjon 7.1.6 på forrige side. Med dette mener man at endring i den ene variabelen kan forklares uttrykt ved lag av differansen mellom tidsrekkene muligens etter skalering og lag av differansen til hver av tidsrekkene. Dette fører oss over til fenomenet Grangerkausalitet. Hvis man har to kointegrerte $I(1)$-variabler, må Grangerkausalitet eksistere i minst en retning. Dvs. at minst en av variablene hjelper å predikere den andre. Dette er beskrevet nærmere i Pfaff (2008, s. 75-78). Vi utfører regresjonen

$$x_{1,t} = \alpha_1 x_{2,t} + z_t \quad \text{for } t = 1, \ldots, T,$$

der $x_{1,t}$ og $x_{2,t}$ begge er $I(1)$. Vi utfører deretter regresjonene:

$$\Delta x_{1,t} = \psi_0 + \gamma_1 \hat{z}_{t-1} + \sum_{i=1}^{K} \psi_{1,i} \Delta x_{2,t-i} + \sum_{i=1}^{L} \psi_{2,i} \Delta x_{1,t-i} + \epsilon_{1,t}$$
$$\Delta x_{2,t} = \xi_0 + \gamma_2 \hat{z}_{t-1} + \sum_{i=1}^{K} \xi_{1,i} \Delta x_{1,t-i} + \sum_{i=1}^{L} \xi_{2,i} \Delta x_{2,t-i} + \epsilon_{2,t}.$$

Dersom vi har kointegrasjon, må $\hat{z}_{t-1}$ ha en signifikant effekt i minst en av regresjonene.

Langtidsprediksjon

En nyttig anvendelse av kointegrasjon er at når man predikerer et stykke frem i tid, vil prediksjonene av de to kointegrerte tidsrekkene forme et konstant forhold. Dette er beskrevet i Zivot og Wang (2006, s. 435).

Vi lar $x_t = (x_{1,t}, \ldots, x_{n,t})'$ uttrykke en vektor med $I(1)$-variabler. Vi har at x_t er kointegrert hvis det eksisterer en $n \times 1$-vektor $\beta = (\beta_1, \ldots, \beta_n)'$ slik at

$$\beta' x_t = \beta_1 x_{1,t} + \ldots + \beta_n x_{n,t} \sim I(0).$$

Den lineære kombinasjonen $\beta' x_t$ kalles ofte langtidslikevektsammenhengen.

Intuisjonen er at $I(1)$-tidsrekker med langtidslikevekt ikke kan bevege seg for langt vekk fra likevekten fordi økonomiske krefter vil forsøke å gjenopprette likevektsammenhengen.

Vi ser på en kointegrasjonsmodell på formen

$$\Delta x_t = \Pi x_{t-1} + \epsilon_t, \tag{7.3}$$

der $x_t = (x_{1,t}, x_{2,t})'$. Dersom x_t er kointegrert eksisterer det en 2×1-vektor $\beta = (\beta_1, \beta_2)'$ slik at

$$\beta' x_t = \beta_1 x_{1,t} + \beta_2 x_{2,t} \sim I(0).$$

Normaliserer vi slik at $\beta_1 = 1$ og $\beta_2 = -\beta$ får vi kointegrasjonssammenhengen

$$\beta' x_t = x_{1,t} - \beta x_{2,t}.$$

Vi får da langtidslikevekten

$$x_{1,t} = \beta x_{2,t} + u_t,$$

der u_t er $I(0)$ og representerer stokastiske avvik fra langtidslikevekten

$$x_{1,t} = \beta x_{2,t}.$$

Siden x_t er kointegrert med én kointegrasjonsvektor har vi at $\text{rang}(\Pi) = 1$ og vi har at

$$\Pi = \alpha\beta' = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \begin{pmatrix} 1 & -\beta \end{pmatrix} = \begin{pmatrix} \alpha_1 & -\alpha_1\beta \\ \alpha_2 & -\alpha_2\beta \end{pmatrix}.$$

Vi kan da skrive (7.3) som

$$\Delta x_t = \alpha\beta' x_{t-1} + \epsilon_t.$$

Dette gir oss at:

$$\Delta x_{1,t} = \alpha_1(x_{1,t-1} - \beta x_{2,t-1}) + \epsilon_{1,t}$$
$$\Delta x_{2,t} = \alpha_2(x_{1,t-1} - \beta x_{2,t-1}) + \epsilon_{2,t}.$$

Hvis vi ser på den første ligningen, forklarer faktoren $x_{1,t-1} - \beta x_{2,t-1}$ hvor langt vekke vi er fra likevektsammenhengen, mens faktoren α_1 forklarer hvor fort vi går mot likevekten. Tilsvarende tolkning får vi for den andre ligningen. Dette er beskrevet i Zivot og Wang (2006, s. 457-458).

Kointegrasjon og høyfrekvens finans

Vi så ovenfor at kointegrasjon ble brukt i økonomi for å se på langtidssammenhenger. I forbindelse med høyfrekvente data har kointegrasjon en noe annen tolkning og er knyttet til arbitrasjeargumenter. Dette er diskutert i Zivot og Wang (2006, s. 436-437).

Vi vil i 7.2 se på triangelarbitrasje. Dersom vi har et triangel av valutakurser, f.eks. USD/SEK, USD/NOK og NOK/SEK, må forholdet mellom de to første være lik den siste. Dersom en av valutakursene avviker for mye fra triangelsammenhengen, vil man ha en arbitrasjemulighet, og økonomiske krefter vil forsøke å eliminere denne arbitrasjemuligheten veldig fort. Å studere triangelarbitrasje med kointegrasjonsmodeller er foreslått av Trapletti *et al.* (2002).

7.2 Tilpasning av kointegrasjonsmodell til valutadata

Som nevnt i 7.1.3 er triangelarbitrasje et kjent fenomen i valutahandel. I Trapletti *et al.* (2002) studerer man data fra valutakursene USD/DEM, USD/JPY og DEM/JPY, mens vi i dette kapittelet studerer data fra valutakursene USD/NOK, USD/SEK og NOK/SEK. Vi har en tidsrekke for hver av valutakursene og tester alle tre tidsrekkene for enhetsrot. Vi estimerer deretter en kointegrasjonsmodell og tester for kointegrasjonsrang. Til slutt forsøker vi å tolke kointegrasjonsmodellen, og vi sammenligner kointegrasjonsmodellens prediksjonsevne med en VAR-modells prediksjonsevne.

7.2.1 Bearbeiding av data

Figuren nedenfor viser hvordan data fra valutahandel i NOK/SEK ser ut. Vi har registrert et tidspunkt samt kjøpers bud(bid) og selgers krav(offer).

time	bid	offer
20080509T002806	1.1777	1.1801
20080509T002912	1.1778	1.1802
20080509T002912	1.1777	1.1801
20080509T002922	1.1776	1.18
20080509T002922	1.1777	1.1801
20080509T002953	1.1778	1.1802
20080509T002953	1.1777	1.1801
20080509T003150	1.1776	1.18
20080509T003150	1.1777	1.1801
20080509T003857	1.1778	1.1802

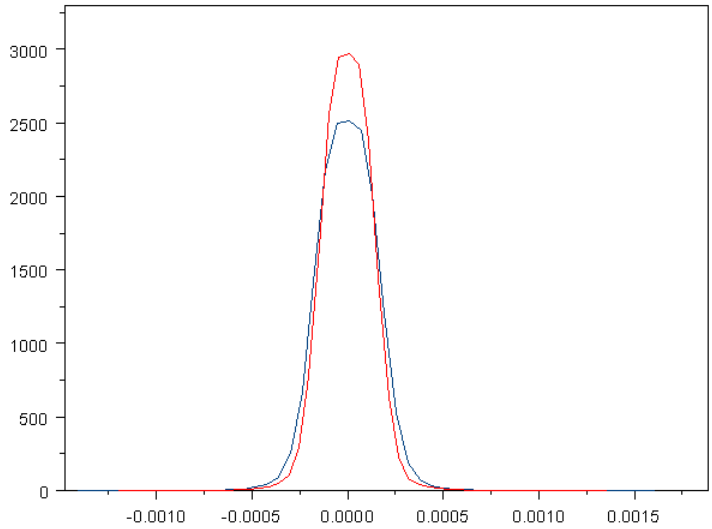
Figur 7.1: Utdrag fra data NOK/SEK 9. mai 2008

Vi har altså ikke oppgitt hva selve prisen ved handelen ble, kun hva kjøper ønsket å betale(bid) og det selger krevde(offer). Vi definerer derfor prisen slik det er gjort i Dacorogna *et al.* (2001, s. 39).

Definisjon 7.2.1: Pris i valutahandel

$$\text{pris} = \frac{\log(\text{bid}) + \log(\text{offer})}{2}$$

Årsaken til at logaritmen til bid og offer inngår, er at dersom man ser på den empiriske fordelingen til de differensierte tidsrekkene, ser man at fordelingen har tykkere haler enn hva som er tilfellet for normalfordelingen. Logaritmefunksjonen demper prisendringene og gjør at fordelingen får smalere haler, noe figuren nedenfor viser.



Figur 7.2: Figuren viser empirisk fordeling til prisendringer NOK/SEK 9. mai 2008 uten bruk av logaritmer(blå) og med bruk av logaritmer(rød)

Handelen i USD/SEK, USD/NOK og NOK/SEK foregår til ulike tidspunktet og med ulik frekvens. For å tilpasse kointegrasjonsmodellen trenger vi tre like lange homogene tidsrekker med observasjoner på de samme tidspunktene. For å løse dette problemet må vi interpolere dataene. Vi har her valgt å bruke en lineær interpolasjon, der vi bruker observasjonen rett før et tidspunkt og observasjonen rett etter tidspunktet. Dette gir oss ligningen

$$z(t_0 + i\Delta t) = z_{j'} + \frac{t_0 + i\Delta t - t_{j'}}{t_{j'+1} - t_{j'}}(z_{j'+1} - z_{j'}),$$

der

$$j' = \max(j | t_j \leq t_0 + i\Delta t),\ t_{j'} \leq t_0 + i\Delta t < t_{j'+1}.$$

Her er i indeksen i den homogene tidsrekken vi får etter interpolasjonen, mens j er indeksen i den opprinnelige inhomogene tidsrekken.

En alternativ måte å interpolere på er å bruke den mest nylige verdien. Vi får da at

$$z(t_0 + i\Delta t) = z_{j'}.$$

Begge metodene er diskutert i Dacorogna *et al.* (2001, s. 37-38). Vi må velge gitteravstanden Δt og har valgt den lik 10 sekunder. I estimeringen av kointegrasjonsmodellen har vi valgt å bruke data fra 9. mai 2008 med de tre valutakursene USD/SEK, USD/NOK og NOK/SEK. Antall observasjoner i de tre tidsrekkene var henholdsvis 14 257, 10 403 og 12 378. Etter interpoleringen har vi i alle tre tidsrekkene 7080 observasjoner.

7.2.2 Test for enhetsrot

Vi ønsker å teste hver av de tre tidsrekkene for enhetsrot, se definisjon 7.1.1 på side 47. Vi har valgt å bruke Augmented Dickey-Fuller enhetsrottest (ADF-test). Denne testen er beskrevet i Pfaff (2008, s. 91-94) og Zivot og Wang (2006, s. 114-121).

Dersom vi har en AR(1)-modell

$$x_t = \phi x_{t-1} + \epsilon_t, \text{ hvor } \{\epsilon_t\} \text{ er iid. og } \epsilon_t \sim \mathcal{N}(0, \sigma^2),$$

kan vi utføre testen:

$$H_0 : \phi = 1, \text{ dvs. enhetsrot } I(1)$$
$$H_1 : |\phi| < 1, \text{ dvs. ikke enhetsrot, men } I(0).$$

Mange finansielle tidsrekker har en struktur som ikke kan fanges opp av en AR(1)-prosess. Vi ønsker derfor en mer generell modell, der vi antar at dataene har ARMA-struktur. Vi utfører regresjonen

$$x_t = c_t + \beta x_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta x_{t-i} + \epsilon_t, \tag{7.4}$$

der c_t er null, en konstant eller $c_t = \omega_0 + \omega_1 t$. Vi kan alternativt utføre den følgende regresjonen som vi har valgt å bruke

$$\Delta x_t = c_t + \pi x_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta x_{t-i} + \epsilon_t, \tag{7.5}$$

der $\pi = \beta - 1$. Vi ønsker da å utføre testen:

$$H_0 : \pi = 0$$
$$H_1 : \pi \neq 0.$$

Testobservatoren er da gitt som

$$\frac{\hat{\pi}}{\widehat{\mathrm{SD}}(\hat{\pi})}.$$

I regresjonen ovenfor må antall lag p velges. Vi velger p slik at siste lag vi tar med er signifikant. Samtidig må p velges så stor at vi ikke har autokorrelasjon i residualene i modellen.

Vi har først testet tidsrekken NOK/SEK for enhetsrot. Som nevnt kan vi ha tre ulike former på c_t i (7.4), og vi får dermed tre ulike former på regresjonen i (7.5):

$$\Delta x_t = \omega_0 + \omega_1 t + \pi x_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta x_{t-i} + \epsilon_t \tag{7.6}$$

$$\Delta x_t = \omega_0 + \pi x_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta x_{t-i} + \epsilon_t \tag{7.7}$$

$$\Delta x_t = \pi x_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta x_{t-i} + \epsilon_t. \tag{7.8}$$

Vi starter med en modell på formen (7.6), og vi kommer frem til at verken ω_0, ω_1 eller π er signifikant ulik null. Vi går derfor videre til modellen på formen (7.7), og vi kommer frem til at verken ω_0 eller π er signifikante. Vi kommer dermed frem til at vi har en modell på formen (7.8), og at vi har en $I(1)$-prosess.

Videre differensierer vi tidsrekken og sjekker den differensierte tidsrekken for enhetsrot. Hensikten med dette er å sjekke om prosessen er $I(2)$, noe vi kommer frem til at den ikke er. Vi konkluderer med at NOK/SEK inneholder en enhetsrot. Ved å utføre de samme testene på tidsrekkene USD/SEK og USD/NOK kommer vi frem til at også disse tidsrekkene inneholder en enhetsrot.

7.2.3 Estimering av kointegrasjonsmodell

Dersom vi kjenner rangen til Π kan vi skrive kointegrasjonsmodellen på formen

$$\Delta x_t = \mu d_t + \alpha\beta' x_{t-1} + \Phi_1^* \Delta x_{t-1} + \ldots + \Phi_{p-1}^* \Delta x_{t-p+1} + \epsilon_t \quad \text{for } t = p+1, \ldots, T.$$

Hvordan vi finner rangen til Π er beskrevet i 7.2.4. Vi antar i fortsettelsen at vi har $\text{rang}(\Pi) = m$.

Vi utfører de to multivariate regresjonene:

$$\Delta x_t = \gamma_0 d_t + \Omega_1 \Delta x_{t-1} + \ldots + \Omega_{p-1} \Delta x_{t-p+1} + u_t$$

$$x_{t-1} = \gamma_1 d_t + \Xi_1 \Delta x_{t-1} + \ldots + \Xi_{p-1} \Delta x_{t-p+1} + v_t.$$

Definerer utvalgskovariansmatrisene:

$$S_{00} = \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{u}_t \hat{u}_t', \; S_{01} = \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{u}_t \hat{v}_t' \text{ og } S_{11} = \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{v}_t \hat{v}_t'.$$

Vi finner deretter egenverdiene og egenvektorene til $S_{10} S_{00}^{-1} S_{01}$ med hensyn til S_{11} ved å løse egenverdiproblemet

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0.$$

Vi uttrykker egenverdiene og egenvektorparene ved $(\hat{\lambda}, e_i)$ hvor $\hat{\lambda}_1 > \hat{\lambda}_2 > \ldots > \hat{\lambda}_k$. Her er egenvektorene normalisert slik at

$$e' S_{11} e = I,$$

hvor $e = [e_1, \ldots, e_k]$ er matrisen med egenvektorer. Det unormaliserte sannsynlighetsmaksimeringsestimatet av kointegrasjonsvektoren β er $\hat{\beta} = [e_1, \ldots e_m]$. Ut fra dette estimatet kan vi finne normalisert β, slik det er forklart i Tsay (2005, s. 383-384). Vi normaliserer slik at første rad i β blir lik 1, og kaller dette $\hat{\beta}_c$. Estimatene av α og de andre parametrene finner vi ved å utføre den multivariate lineære regresjonen

$$\Delta x_t = \mu d_t + \alpha \hat{\beta}_c' x_{t-1} + \Phi_1^* \Delta x_{t-1} + \ldots + \Phi_{p-1}^* \Delta x_{t-p+1} + \epsilon_t \quad \text{for } t = p+1, \ldots T.$$

7.2.4 Test for kointegrasjonsrang

Kointegrasjonsmodellen er gitt som

$$\Delta x_t = \mu d_t + \Pi x_{t-1} + \Phi_1^* \Delta x_{t-1} + \ldots + \Phi_{p-1}^* \Delta x_{t-p+1} + \epsilon_t \quad \text{for } t = p+1, \ldots, T.$$

Vi ønsker å teste rangen til Π, dvs. antall ikke-null egenverdier til Π. Her er Π relatert til kovariansmatrisen mellom x_{t-1} og Δx_t etter å ha justert for effekten av d_t og Δx_{t-i} for $i = 1, \ldots, p-1$. De justerte tidsrekkene for x_{t-i} og Δx_t er henholdsvis $\hat{v}_t$ og $\hat{u}_t$, hvor $\hat{v}_t$ og $\hat{u}_t$ er som i forrige avsnitt. Ligningen av interesse blir da:

$$\hat{u}_t = \Pi \hat{v}_t + \epsilon_t.$$

Vi vil se på to tester, og begge er basert på de estimerte egenverdiene $\hat{\lambda}_1 > \hat{\lambda}_2 > \ldots > \hat{\lambda}_k$ til matrisen Π. Testene er beskrevet i Tsay (2005, s. 384-385).

Tracekointegrasjonstest

Vi ønsker å teste:

$$H_0 : \text{Rang}(\Pi) = m$$
$$H_1 : \text{Rang}(\Pi) > m.$$

Observatoren er gitt som

$$\text{LK}_{\text{tr}}(m) = -(T-p) \sum_{i=m+1}^{k} \log(1 - \hat{\lambda}_i).$$

Hvis $\text{Rang}(\Pi) = m$, vil $\hat{\lambda}_i$ være liten for $i > m$ og dermed vil $\text{LK}_{\text{tr}}(m)$ være liten. Kritiske verdier må man finne ved simulering. Nedenfor ser vi resultatene ved test for kointegrasjon. Testen gir en sterk indikasjon på at vi har kointegrasjonsrang lik én, ettersom vi ikke forkaster H_0 om $r \leq 1$, men forkaster H_0 om $r = 0$.

	Observator	10 %	5 %	1 %
$r \leq 2$	0,88	6,50	8,18	11,65
$r \leq 1$	5,48	15,66	17,95	23,52
$r = 0$	819,17	28,71	31,52	37,22

Tabell 7.1: Trace-observator med kritiske verdier

Maksimumegenverdi-test

Man kan også teste antall kointegrasjonsvektorer ved å utføre testen:

$$H_0 : \text{Rang}(\Pi) = m$$
$$H_1 : \text{Rang}(\Pi) = m + 1.$$

Observatoren er gitt som

$$\text{LK}_{\max}(m) = -(T - p)\log(1 - \hat{\lambda}_{m+1}).$$

Også her må man finne de kritiske grensene ved simulering, i følge Tsay (2005, s. 384-385). Vi ser nedenfor at maksimumegenverdi-testen gir samme konklusjon som tracekointegrasjonstesten.

	Observator	10 %	5 %	1 %
$r \leq 2$	0,88	6,50	8,18	11,65
$r \leq 1$	4,60	12,91	14,90	19,19
$r = 0$	813,69	18,90	21,07	25,75

Tabell 7.2: Maksimumegenverdi-observator med kritiske verdier

7.2.5 Økonomisk tolkning av modell

Vi estimerer modellen som beskrevet i 7.2.3, med $\text{rang}(\Pi) = 1$, som vi kom frem til 7.2.4. Vi får da at de estimerte verdiene for β og α er som følger

$$\hat{\beta}_c = \begin{pmatrix} 1{,}000 \\ -1{,}001 \\ 1{,}002 \end{pmatrix} \text{ og } \hat{\alpha} = \begin{pmatrix} -0{,}004 \\ 0{,}182 \\ -0{,}265 \end{pmatrix}.$$

Tolkningen av $\hat{\alpha}$ er beskrevet i 7.1.3, og vi så der at α-verdiene forteller oss hvor fort man når likevekten. Ut fra de estimerte verdiene ser vi at USD/NOK med α-verdi lik -0,004 ser ut til å nå likevekten senere enn de to andre valutakursene.

7.2.6 Test av kointegrasjonsmodellens prediksjonsevne

Vi ønsker å teste kointegrasjonsmodellens evne til å predikere retningen til valutakursene. For å teste dette har vi valgt å sammenligne kointegrasjonsmodellen og VAR-modellens prediksjonsevne. Vi ser først på hvordan kointegrasjonsmodellen kan brukes til prediksjon.

Prediksjon av kointegrasjonsmodell

Kointegrasjonsmodellen er gitt som

$$\Delta x_t = \mu d_t + \Pi x_{t-1} + \Phi_1^* \Delta x_{t-1} + \ldots + \Phi_{p-1}^* \Delta x_{t-p+1} + \epsilon_t \quad \text{for } t = p+1, \ldots T.$$

Vi kan benytte dette for å predikere Δx_t. Ettstegsprediksjon blir:

$$\Delta x_t(1) = \mu d_t + \Pi x_t + \Phi_1^* \Delta x_t + \ldots + \Phi_{p-1}^* \Delta x_{t-p+2},$$

med feil $e_t(1) = \epsilon_{t+1}$ som har kovariansmatrise Σ. Mens tostegsprediksjonen blir:

$$\begin{aligned}\Delta x_t(2) =& \mu d_t + \Pi x_{t+1} + \Phi_1^* \Delta x_{t+1} + \Phi_2^* \Delta x_t + \ldots + \Phi_{p-1}^* \Delta x_{t-p+3} \\ =& \mu d_t + \Pi(x_t + \Delta x_t(1)) + \Phi_1^* \Delta x_t(1) + \Phi_2^* \Delta x_t + \ldots + \Phi_{p-1}^* \Delta x_{t-p+3},\end{aligned}$$

med feil lik

$$\begin{aligned}e_t(2) =& \epsilon_{t+2} + \Pi(x_{t+1} - (x_t + \Delta x_t(1))) + \Phi_1^*(\Delta x_{t+1} - \Delta x_t(1)) \\ =& \epsilon_{t+2} + (\Pi + \Phi_1^*)\epsilon_{t+1},\end{aligned}$$

som har kovariansmatrise $\Sigma + (\Pi + \Phi_1^*)\Sigma(\Pi + \Phi_1^*)'$.

Vi ser at vi kan predikere videre fremover ved å regne ut prediksjonene rekursivt. Denne fremgangsmåten er beskrevet i Zivot og Wang (2006, s. 398-400).

Prediksjonsforsøk

I prediksjonsforsøket sammenligner vi en kointegrasjonsmodell på formen (7.2) med fem lag, med en VAR-modell for Δx_t med samme antall lag. Forskjellen mellom de to modellene er at leddet Πx_{t-1} ikke inngår i VAR-modellen, dvs. kointegrasjonssammenhengen. Vi ønsker å vise at man får bedre prediksjoner ved å utnytte at man har kointegrasjon.

Oppsettet for forsøket er som følger:
1) Tilpass VAR(5)-modell for Δx_t, se (7.1), og kointegrasjonsmodell, se (7.2).
2) Prediker Δx_t $1, 2, \ldots,10$ steg fremover i hver av de to modellene.
3) Sjekker om

$$\text{sgn}(\Delta x_{i,t}(k)) = \text{sgn}(\Delta x_{i,t+k}) \quad \text{for } i = 1, 2, 3 \text{ og } k = 1, \ldots,10,$$

dvs. om vi har predikert riktig retning på $\Delta x_{i,t}$ i hver av de to modellene.

Vi gjentar de tre stegene ved å bevege oss fremover i tid. Etterhvert som vi får flere observasjoner, oppdaterer vi også estimeringen av VAR(5)-modellen og kointegrasjonsmodellen. Vi har i forsøket gjentatt de tre stegene 3500 ganger. Til slutt regner vi ut andel riktige prediksjoner av retning. Resultatene ser vi i tabellene på neste side, som viser andel riktige prediksjoner av retning i 1-10 stegsprediksjon i hver av valutakursene. Vi ser at kointegrasjonsmodellen ser ut til å gjøre det bedre enn VAR-modellen, særlig ved prediksjoner et stykke fremover. Vi ser også at kointegrasjonsmodellen har mest problemer med å predikere valutakursen USD/NOK. Dette er i tråd med funnene i 7.2.5 der vi så at USD/NOK beveget seg senere mot likevekten enn de to andre valutakursene.

	1	2	3	4	5	6	7	8	9	10
USD/NOK	0,46	0,48	0,50	0,49	0,49	0,50	0,49	0,50	0,48	0,49
USD/SEK	0,53	0,50	0,49	0,49	0,51	0,50	0,49	0,49	0,50	0,49
NOK/SEK	0,51	0,49	0,50	0,49	0,49	0,48	0,49	0,51	0,50	0,50

Tabell 7.3: Andel riktige i 1-10 stegsprediksjon VAR-modell

	1	2	3	4	5	6	7	8	9	10
USD/NOK	0,46	0,49	0,51	0,52	0,51	0,52	0,52	0,52	0,51	0,51
USD/SEK	0,54	0,55	0,55	0,55	0,55	0,55	0,55	0,54	0,54	0,54
NOK/SEK	0,52	0,55	0,56	0,57	0,58	0,57	0,56	0,58	0,57	0,58

Tabell 7.4: Andel riktige i 1-10 stegsprediksjon kointegrasjonsmodell

7.3 Videre studier av modellen

Et tema som ikke er tatt opp her, er om det virkelig er mulig å bruke kointegrasjonsmodellen til å tjene penger med. Det er viktig å merke seg at dersom man vil finne ut dette, må man dra inn transaksjonskostnader og forskjell i likviditet i de ulike valutakursene.

Litteratur

Ait-Sahalia Y., Mykland P.A. og Zhang L. (2005). «Ultra high frequency volatility estimation with dependent microstructure noise». *Discussion Paper Series 1: Economic Studies 2005,30*, Deutsche Bundesbank, Research Centre. URL: `http://ideas.repec.org/p/zbw/bubdp1/4224.html`. Referert til på side 44.

Bauwens L. og Giot P. (2000). «The logarithmic acd model: An application to the bid-ask quote process of three nyse stocks». *Annales d'Économie et de Statistique*, , nummer 60, side 117–149. ISSN 0769489X. URL: `http://www.jstor.org/stable/20076257`. Referert til på side 22.

Bauwens L., Pohlmeier W. og Veredas D. (red.) (2008). *High Frequency Financial Econometrics*. Physica-Verlag. Referert til på side 7, 38 og 43.

Berndt E., Hall B., Hall R. og Hausman J. (1974). «Estimation and inference in nonlinear structural models». *Annals of Economic and Social Measurement*, **volum 3**, nummer 4, side 103–116. Referert til på side 32.

Bessembinder H. (2003). «Trade execution costs and market quality after decimalization». *The Journal of Financial and Quantitative Analysis*, **volum 38**, nummer 4, side 747–777. ISSN 00221090. DOI: 10.2307/4126742. Referert til på side 11 og 12.

Dacorogna M.M., Gençay R., Müller U.A., Olsen R.B. og Pictet O.V. (2001). *An Introduction to High-Frequency Finance*. Academic Press. Referert til på side 53 og 54.

Dobson A.J. (2002). *An introduction to generalized linear models*. Chapman & Hall/CRC. Referert til på side 23, 29, 42, 62 og 63.

Engle R.F. og Russell J.R. (1998). «Autoregressive conditional duration: A new model for irregularly spaced transaction data». *Econometrica*, **volum 66**, nummer 5, side 1127–1162. DOI: 10.2307/2999632. Referert til på side 21.

Gerald C.F. og Wheatley P.O. (2004). *Applied Numerical Analysis*. Addison-Wesley. Referert til på side 33 og 36.

Granger C.W. (2003). «Time series analysis, cointegration and applications (the nobel lecture)». URL: `http://nobelprize.org/nobel_prizes/economics/laureates/2003/granger-lecture.pdf`. Nobel Lecture, December 2003. Referert til på side 50.

Grøtte O. (2006). *Aksjekjøp og daytrading - Metode, psykologi, risiko og strategier*. Hegnar Media. Referert til på side 10, 11 og 12.

Hendershott T. og Moulton P.C. (2007). «The shrinking new york stock exchange floor and the hybrid market». Referert til på side 12.

NYSE (2006). *A Guide to the NYSE Markeplace*. URL: `http://www.nyse.com/pdfs/nyse_bluebook.pdf`. Referert til på side 12.

O'Hara M. (1995). *Market Microstructure Theory*. WileyBlackwell. Referert til på side 41 og 42.

OSE (2006). *Facts & Figures Oslo Børs Stock Market 2006*. Referert til på side 10.

OSE (2007a). *Aksjer for alle - en kort innføring for deg som vil vite mer om aksjemarkedet*. Referert til på side 9.

OSE (2007b). *Guide to the Norwegian stock market 2007*. Referert til på side 9, 10 og 11.

Pfaff B. (2008). *Analysis of Integrated and Cointegrated Time Series with R*. Springer. Referert til på side 8, 47, 48, 50 og 54.

Rasmus S. (2007). *Derivative Pricing*. Lund Institute of Technology. Referert til på side 46.

Russell J.R. og Engle R.F. (2005). «A discrete-state continuous-time model of financial transactions prices and times: The autoregressive conditional multinomial-autoregressive conditional duration model». *Journal of Business & Economic Statistics*, **volum 23**, side 166–180. DOI: 10.1198/073500104000000541. Referert til på side 7, 8, 20, 21, 25, 26, 30, 38, 43 og 44.

Sandvik B. (2003). *Innføring i finansteori*. Fagbokforlaget. Referert til på side 42.

Sørensen B.E. (2005). «Cointegration». URL: `http://141.217.212.112/coint.pdf`. Referert til på side 48.

Taylor H.M. og Karlin S. (1998). *An Introduction to Stochastic Modeling*. Academic Press. Referert til på side 44.

Trapletti A., Geyer A. og Leisch F. (2002). «Forecasting exchange rates using cointegration models and intra-day data». *Journal of Forecasting*, **volum 21**, side 151–166. DOI: 10.1002/for.822. Referert til på side 8 og 52.

Tsay R.S. (2005). *Analysis of Financial Time Series*. Wiley-Interscience. Referert til på side 15, 16, 18, 21, 36, 47, 48, 56 og 57.

Walpole R.E., Myers R.H., Myers S.L. og Ye K. (2002). *Probability & Statistics For Engineers & Scientists*. Prentice Hall. Referert til på side 19.

Zivot E. og Wang J. (2006). *Modeling Financial Time Series with S-PLUS*. Springer Science+Business Media, Inc. Referert til på side 50, 51, 52, 54 og 58.

A
Logistisk regresjon

I 4.2 utfører vi i (4.11) den nominale logistiske regresjonen

$$\log\left(\frac{\pi_i}{1-\iota'\pi_i}\right) = c + \sum_{j=1}^{p} A_j(x_{i-j} - \pi_{i-j}) + \sum_{j=1}^{q} B_j h(\pi_{i-j}) + \chi \log((\tau_i, \tau_{i-1})'). \quad \text{(A.1)}$$

For å forstå denne modellen ser vi først på logistisk regresjon med binær responsvariabel, slik det er fremstilt i Dobson (2002, s. 115-121).

A.1 Logistisk regresjon med binær responsvariabel

Vi starter med å definere en binær stokastisk variabel.

Definisjon A.1.1: Binær stokastisk variabel

$$q = \begin{cases} 1 & \text{hvis utfallet er suksess} \\ 0 & \text{hvis utfallet er feil,} \end{cases}$$

med sannsynligheter $\mathbb{P}(q = 1) = \pi$ og $\mathbb{P}(q = 0) = 1 - \pi$.

Vi tenker oss at vi har n slike stokastiske variabler $q_1, \ldots, q_n$, der $\mathbb{P}(q_i = 1) = \pi_i$. Vi ønsker å modellere q_i og ser først på modellen

$$\pi_i = z_i'\beta,$$

der z_i her er en kolonnevektor med forklaringsvariabler. Ulempen med denne modellen er at π_i ikke er begrenset til området $0 \leq \pi \leq 1$. Vi ønsker oss derfor en modell som sikrer oss at dette kravet er oppfylt. For å sikre dette modellerer man ofte ved å bruke en kumulativ fordelingsfunksjon

$$\pi = \int_{-\infty}^{t} f(s)\,\mathrm{d}s,$$

hvor $f(s) \geq 0$ og $\int_{-\infty}^{\infty} f(s)\,\mathrm{d}s = 1$. Sannsynlighetstettheten $f(s)$ kalles toleransefunksjonen. Et valg av toleransefunksjon er

$$f(s) = \frac{\beta_2 \exp(\beta_1 + \beta_2 s)}{[1 + \exp(\beta_1 + \beta_2 s)]^2}$$

så

$$\pi_i = \int_{-\infty}^{z_i} f(s)\,\mathrm{d}s = \frac{\exp(\beta_1 + \beta_2 z_i)}{1 + \exp(\beta_1 + \beta_2 z_i)}.$$

Dette gir lenkefunksjonen

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_1 + \beta_2 z_i.$$

Funksjonen $\log\left(\frac{\pi}{1-\pi}\right)$ kalles logit-funksjonen, og en modell på denne formen kalles en logistisk modell. Den logistiske modellen kan generaliseres til en modell på formen

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = z_i'\beta,$$

der z_i og β er kolonnevektorer.

A.2 Nominal logistisk regresjon

I A.1 har vi en variabel som har to kategorier, og vi ønsker å generalisere til tilfellet med flere enn to kategorier, og vi følger tilnærmingen i Dobson (2002, s. 135-137). Vi betrakter en stokastisk variabel q med J kategorier, der $\pi_1, \ldots, \pi_J$ uttrykker sannsynlighetene med $\pi_1 + \ldots + \pi_J = 1$.
Nominal logistisk regresjonsmodeller brukes dersom det ikke er noen naturlig ordning blant responsvariablene. En kategori velges vilkårlig som referansekategori. Anta at dette er den første kategorien. Da er logit for de andre kategoriene gitt som

$$\text{logit}(\pi_j) = \log\left(\frac{\pi_j}{\pi_1}\right) = z_j'\beta_j \quad \text{for } j = 2, \ldots, J. \tag{A.2}$$

De $J-1$ logit-ligningene brukes samtidig for å estimere vektorene med parametre β_j, $j = 2, \ldots, J$. Når parameterestimatet b_j er funnet, kan de lineære prediktorene $z_j'b_j$ regnes ut. Fra (A.2) får vi at

$$\hat{\pi}_j = \hat{\pi}_1 \exp(z_j'b_j) \quad \text{for } j = 2, \ldots, J.$$

Men $\hat{\pi}_1 + \ldots + \hat{\pi}_J = 1$, så

$$\hat{\pi}_1 = \frac{1}{1 + \sum_{j=2}^{J} \exp(z_j'b_j)}$$

$$\hat{\pi}_j = \frac{\exp(z_j' b_j)}{1 + \sum_{j=2}^{J} \exp(z_j' b_j)} \quad \text{for } j = 2, \ldots, J.$$

A.3 Nominal logistisk regresjon i ACM-modellen

Vi ønsker nå å relatere teorien om nominal logistisk regresjon til ACM-modellen. I (4.7) har vi at

$$g(y_i | y_{1:i-1}, \tau_{1:i}) = \begin{cases} \pi_{i,-1} & y_i < 0 \\ \pi_{i,0} & y_i = 0 \\ \pi_{i,1} & y_i > 0, \end{cases}$$

og vi har dermed tre kategorier. Vi ser at sannsynlighetene avhenger av indeksen i, og vi får en modell på formen

$$\text{logit}(\pi_{i,j}) = \log\left(\frac{\pi_{i,j}}{\pi_{i,0}}\right) = z_{i,j}' \beta_j \quad \text{for } j = -1, 1 \tag{A.3}$$

og

$$\hat{\pi}_{i,j} = \frac{\exp(z_{i,j}' b_j)}{1 + \sum_{j=-1,1} \exp(z_{i,j}' b_j)} \quad \text{for } j = -1, 1$$
$$\hat{\pi}_{i,0} = 1 - (\hat{\pi}_{i,-1} + \hat{\pi}_{i,1}).$$

Eksempel A.3.1

Vi ønsker å vise ved et konkret eksempel at (A.1) kan skrives på formen (A.3). Vi setter $p = 1$ og $q = 1$ og forenkler modellen noe. Vi får da modellen

$$\log\left(\frac{\pi_i}{1 - \iota' \pi_i}\right) = c + A_1(x_{i-1} - \pi_{i-1}) + B_1 h(\pi_{i-1})$$
$$\begin{bmatrix} \text{logit}(\pi_{i,-1}) \\ \text{logit}(\pi_{i,1}) \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \epsilon_{i-1,-1} \\ \epsilon_{i-1,1} \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ 0 & b_{22} \end{bmatrix} \begin{bmatrix} h(\pi_{i-1,-1}) \\ h(\pi_{i-1,1}) \end{bmatrix},$$

der $\epsilon_{i,j} = x_{i,j} - \pi_{i,j}$. Det første elementet i vektoren på venstre side kan skrives som

$$\text{logit}(\pi_{i,-1}) = z_{i,-1}' \beta_{-1},$$

der

$$z_{i,-1}' = \begin{bmatrix} 1 & \epsilon_{i-1,-1} & \epsilon_{i-1,1} & h(\pi_{i-1,-1}) \end{bmatrix} \text{ og}$$
$$\beta_{-1}' = \begin{bmatrix} c_1 & a_{11} & a_{12} & b_{11} \end{bmatrix}.$$

B
Programkode

B.1 MATLAB-kode for simulering av ACM-ACD-modellen

```
function[y,tau] = ACMACDsim(antall,c,A1,A2,A3,B1,B2,B3,kji,omega,alpha1,
    alpha2,alpha3,beta1,beta2,beta3)
    %function[y,tau] = ACMACDsim(antall,c,A1,A2,A3,B1,B2,B3,kji,omega,
        alpha1,alpha2,alpha3,beta1,beta2,beta3)
    % c 2x1-matrise, A1,A2,A3,B1,B2,B3,kji 2x2-matriser, omega,alpha1,
        alpha2,alpha3,beta1,beta2,beta3 skalarverdier
    if(beta1+beta2+beta3>1)
        disp('Velg␣beta1+beta2+beta3<1');
    end
    if(min(abs(roots([B3(1,1),B2(1,1),B1(1,1),-1])))<1);
        disp('Alle␣z␣som␣tilfredstiller␣|I-B1z-B2z^2-B3z^3|=0␣ligger␣
            ikke␣utenfor␣enhetssirkelen');
    end
    envek = [1 1];
    psi(1:antall) = 1;
    tau(1:antall) = 1;
    y(1:antall) = 0;
    x = zeros(2,antall);
    pi = zeros(2,antall) + 0.1;
    for i=4:antall
        psi(i) = exp(omega + alpha1*(tau(i-1)/psi(i-1)) + alpha2*(tau(i
            -2)/psi(i-2)) + alpha3*(tau(i-3)/psi(i-3)) + beta1*log(psi(i
            -1)) + beta2*log(psi(i-2)) + beta3*log(psi(i-3)));
        tau(i) = exprnd(psi(i),1,1);
        midl = exp( c + A1*(x(1:2,i-1)-pi(1:2,i-1)) + + A2*(x(1:2,i-2)-
            pi(1:2,i-2)) + A3*(x(1:2,i-3)-pi(1:2,i-3)) + B1*log(pi(1:2,i
            -1)/(1-(envek*pi(1:2,i-1)))) + B2*log(pi(1:2,i-2)/(1-(envek*
```

```
            pi(1:2,i-2))) + B3*log(pi(1:2,i-3)/(1-(envek*pi(1:2,i-3))))
            + kji*[log(tau(i)); log(tau(i-1))] );
        pi(1:2,i) = midl/(1+(envek*midl));
        sim = mnrnd(1,[pi(1,i) (1-(pi(1,i)+pi(2,i))) pi(2,i)]);
        if(sim(1)==1)
            y(i) = -0.01; x(1:2,i)=[1;0];
        elseif(sim(2)==1)
            y(i) = 0; x(1:2,i)=[0;0];
        else
            y(i) = 0.01; x(1:2,i)=[0;1];
        end
    end
```

B.2 R-kode for estimering av ACM-ACD-modellen

```
ACMest <- function(y, tau, pqlag, startverdi) {
# y og tau vektorer med henholdsvis presendringer og tidsavstander
# pqlag settes lik 1,2 eller 3 for estimering av henholdsvis ACM(1,1),
    ACM(2,2) eller ACM(3,3)-modell
# startverdi=c(c1,kji11,kji12,kji21,kji22,a1_11,a1_21,b1_11,a2_11,a2_21,
    b2_11,a3_11,a3_21,b3_11)
    if (pqlag < 1 || pqlag > 3) {
        stop(paste("Velg␣pqlag␣lik␣1,2␣eller␣3"))
    }
    if (length(startverdi) < (5 + pqlag * 3)) {
        stop(paste("Alle␣startverdiene␣er␣ikke␣spesifisert"))
    }
    if (pqlag == 1) {
        rotter = abs(polyroot(c(-1, startverdi[8])))
    }
    else if (pqlag == 2) {
        rotter = abs(polyroot(c(-1, startverdi[8], startverdi[11])))
    }
    else if (pqlag == 3) {
        rotter = abs(polyroot(c(-1, startverdi[8], startverdi[11],
            startverdi[14])))
    }
    if (min(rotter) < 1) {
        stop(paste("Alle␣z␣som␣tilfredstiller␣|I-B1z-B2z^2-B3z^3|=0␣
            ligger␣ikke␣utenfor␣enhetssirkelen"))
    }
    antall <<- length(y)
    x <- matrix(0, 2, antall)
    x3 <- matrix(0, 3, antall)
    for (i in 1:antall) {
        if (y[i] < 0) {
```

```
            x[, i] <- matrix(c(1, 0), 2, 1)
            x3[, i] <- matrix(c(1, 0, 0), 3, 1)
        }
        else if (y[i] == 0) {
            x[, i] <- matrix(c(0, 0), 2, 1)
            x3[, i] <- matrix(c(0, 1, 0), 3, 1)
        }
        else if (y[i] > 0) {
            x[, i] <- matrix(c(0, 1), 2, 1)
            x3[, i] <- matrix(c(0, 0, 1), 3, 1)
        }
    }
    tau <<- tau
    pqlag <<- pqlag
    x <<- x
    x3 <<- x3
    library(maxLik)
    maxBHHH(likelihooden, start = startverdi, print.level = 0,
        iterlim = 300)
}
likelihooden <- function(param) {
    c <- matrix(c(param[1], param[1]), 2, 1)
    kji <- matrix(c(param[2], param[4], param[3], param[5]),
        2, 2)
    A1 <- matrix(c(param[6], param[7], param[7], param[6]),
        2, 2)
    B1 <- matrix(c(param[8], 0, 0, param[8]), 2, 2)
    A2 <- mat.or.vec(2, 2)
    B2 <- mat.or.vec(2, 2)
    A3 <- mat.or.vec(2, 2)
    B3 <- mat.or.vec(2, 2)
    if (pqlag >= 2) {
        A2 <- matrix(c(param[9], param[10], param[10], param[9]),
            2, 2)
        B2 <- matrix(c(param[11], 0, 0, param[11]), 2, 2)
    }
    if (pqlag == 3) {
        A3 <- matrix(c(param[12], param[13], param[13], param[12]),
            2, 2)
        B3 <- matrix(c(param[14], 0, 0, param[14]), 2, 2)
    }
    envek <- matrix(c(1, 1), 1, 2)
    pi <- matrix(0.1, 2, antall)
    hpi <- matrix(log(0.125), 2, antall)
    funksjonen <- matrix(0, antall, 1)
    for (i in 4:antall) {
        hpi[, i] <- c + A1 %*% (x[, i - 1] - pi[, i - 1]) +
            A2 %*% (x[, i - 2] - pi[, i - 2]) + A3 %*% (x[,
```

```
            i - 3] - pi[, i - 3]) + B1 %*% hpi[, i - 1] +
            B2 %*% hpi[, i - 2] + B3 %*% hpi[, i - 3] + kji %*%
            matrix(c(log(tau[i]), log(tau[i - 1])), 2, 1)
        pi[, i] <- exp(hpi[, i])/(1 + envek %*% exp(hpi[,
            i]))
        funksjonen[i] <- t(x3[, i]) %*% log(c(pi[1, i], (1 -
            pi[1, i] - pi[2, i]), pi[2, i]))
    }
    funksjonen
}
```

```
ACDest <- function(tau, plag, qlag, startverdi) {
# tau vektor med tidsavstander
# startverdi=c(omega,alpha1,alpha2,...,beta1,beta2,...)
    tau <<- tau
    plag <<- plag
    qlag <<- qlag
    if (plag < 1 || qlag < 1) {
        stop(paste("Velg␣plag>0␣og␣qlag>0"))
    }
    if (length(startverdi) != (plag + qlag + 1)) {
        stop(paste((1 + plag + qlag), "startverdier␣skal␣spesifiseres"))
    }
    if (sum(startverdi[(plag + 2):length(startverdi)]) >
        1) {
        stop(paste("Velg␣beta1+beta2+...<1"))
    }
    library(maxLik)
    maxBHHH(likelihooden, start = startverdi, print.level = 0,
        iterlim = 300)
}
likelihooden <- function(param) {
    antall <- length(tau)
    omega <- param[1]
    alpha <- param[2:(plag + 1)]
    beta <- param[(plag + 2):length(param)]
    psi <- matrix(1, 1, antall)
    funksjonen <- matrix(0, antall, 1)
    for (i in (max(plag, qlag) + 1):antall) {
        psi[i] <- exp(omega + (tau[(i - 1):(i - plag)]/psi[(i -
            1):(i - qlag)]) %*% alpha + log(psi[(i - 1):(i -
            qlag)]) %*% beta)
        funksjonen[i] <- log((1/psi[i]) * exp(-tau[i]/psi[i]))
    }
    funksjonen
}
```